USING AN OPEN HYPERMEDIA SYSTEM TO DEVELOP NEW TECHNIQUES IN ADAPTIVE HYPERMEDIA

By

Jatinder Hothi

PhD Thesis

University Of Southampton

Electronics and Computer Science Department

30th April 2001

Abstract

The research described in this thesis presents an investigation and evaluation into the field of adaptive hypermedia techniques. The aim of this research was to develop, implement and test a new adaptive hypermedia technique i.e. shading fragments of text. The course of research started off by carrying out a study to develop a simple student model framework containing both static and dynamic variables that could be used for educational adaptive applications. The first empirical study carried out compared three hypermedia applications based on dating techniques in archaeology, one using adaptive navigation support, one using shading to grey out complex fragments of text. These two adaptive applications were compared against an existing (non-adaptive) Archaeology application that had been used in the Department of Archaeology for the previous three years. An evaluation study was carried out to compare the usability aspects of these applications using pre and post-session questionnaires and a log of interaction was also recorded. This study found that shading fragments of text was an accepted method of adaptive hypermedia. It also showed that some of the static variables previously identified as part of the student model such as educational background were very important.

The second empirical study also compared the usability of three applications based on the same content i.e. dating techniques in archaeology. One supported the shading method of adaptive hypermedia and adaptive navigational support in the form of dynamic link annotation by changing the colour of links depending on their state i.e. ready-to-be-learned, or not-ready-to-be-learned. All visited links were underlined. The second application only supported dynamic link annotation without shading and the third was again the existing (non-adapted) hypermedia application was used as a control application. Again, the results reflected user attitudes towards the shading from the first study i.e. that it was a useful adaptive hypermedia technique especially when combined with link annotation. In general users also performed better when using this application, for example they re-visited less documents and visited more glossary terms. This study also revealed that time could be used as a variable in a dynamic student model to predict and assess user performance.

Contents Page

Table of Figures

Table of Tables

Preface

This study was undertaken due to my interest in educational hypermedia. Subsequently my interest and excitement has moved on into the area of adaptive hypermedia techniques and evaluation methods.

Acknowledgements

Above all I thank my supervisor, Professor Wendy Hall CBE, who has encouraged me since my first day, Tim Sly from the Department of Archaeology who provided us with the existing Archaeology Application and was a great help during the evaluation stages and Dr. Ian Heath who kindly developed the Microcosm filter for the second experiment. I would also like to thank my family; my husband, Pinder who has supported me throughout my studies, my daughter Saffron, who is the most wonderful, independent and free spirited child and my parents who have also encouraged me to keep going and where I dumped my daughter when I needed time to work and finally my husband's parents. Also all my friends from the Electronics and Computer Science Department who have seen me through the write up stage, especially Samhaa El-Beltagy who helped me to remain calm each time I was in a panic.

I would also like to thank the Interactive Learning Centre at the University of Southampton who funded my first two years and the Electronics and Computer Science Department who funded the last.

Chapter 1 Introduction

1.1 Introduction

Since it's inception, possibly as early as 1945 (Bush, 1945), researchers in the field of hypertext have strived to drive forward the theory and ideas behind hypertext. Over the years this has resulted in the use of a combination of media, navigation techniques, varying document structures and increasingly user-friendly interfaces. It was Ted Nelson (1980) who coined the terms hypertext and hypermedia around the 1960s. Hypertext refers to the associations between non-linear text. Hypermedia refers to the associations between mixed media types i.e. graphics, video etc. The following quote summarises hypermedia (Lowe et al., 1998).

> An application that uses associative relationships among information contained within multiple media data for the purpose of facilitating access to, and manipulation of, the information encapsulated by the data.

The terms hypertext and hypermedia are often used interchangeably. Today, despite the mass acceptance of the world of the World Wide Web, problems with accepting such technology still exist which prevent the uptake of hypermedia in areas such as education, which can be one of the most useful applications of hypermedia technology. The possible advantages of hypermedia in education are obvious, for example one to one tuition. However, putting what we know into practice has proved to be more of a challenge due to many factors. Some still feel that educational hypermedia applications can only be useful as a reference tool and cannot replace a human teacher as they lack knowledge of the student, the ability to communicate with and guide the student, i.e. all attributes which human teachers bring into the classroom.

Traditional educational systems such as computer based learning (CBL) and computer assisted learning (CAL) either direct the course of learning by controlling the dialogue between the machine and the user, or are exploratory systems that encourage the user to explore concepts using a constrained tool set. However, hypertext tends not to lie within either of these categories as they do not support the freedom to browse the available links in any order which is the philosophy behind hypertext (Hall, 2000).

Adaptive hypermedia (AH) is a spin off from research into hypermedia. AH can be viewed as the next stage in the evolution of hypermedia within the educational paradigm. Essentially, AH represents the integration of artificial intelligence techniques with hypermedia. Researchers in the AH community believe that they are on the edge of a breakthrough in educational teaching/ learning technology, since the technology exists and now the theory can be applied with the implementation of adaptive educational hypermedia applications. In effect, they argue that the use of AH can increase the functionality of hypermedia. Hypermedia applications have already broken through traditional educational barriers and AH applications can take educational computing one step further by making the virtual learning environment (the application) more comfortable, accessible and personal to individual users. This is achieved by tailoring the environment to specific user needs and requirements as opposed to conventional CAL material, which in general presents the same documents, links, and text to all types of users.

Technology within the educational environment must be utilised and presented in a user-friendly manner allowing all levels of users access to the technology, including users without previous experience. AH can ease this process so that users feel they are conversing with an intelligent machine, which would guide them just like a human teacher, not to necessarily replace human teachers but to complement and support the learning process.

Many AH techniques exist which can manipulate the links and/ or the content of the hypertext. Because AH is such a new research area, little work has been done to show which techniques are the most effective. Evaluation within AH applications is also an area, which has not been widely researched. Working with adaptivity, by definition makes evaluating such applications a difficult and complex task.

The research described in this thesis started by looking into the construction of a basic student model (SM) framework for use in educational adaptive applications. Due to the complexity of the resulting SM called the SaDy (Static and Dynamic) SM, it was necessary to test the model. To do this it was decided to carry out an empirical investigation to gather data to populate and test the SaDy SM. This investigation resulted in a simultaneous investigation into adaptive hypermedia techniques, which steered the research into adaptive hypermedia as opposed to research into student modelling per se.

It is difficult to move AH research on, present realistic solutions to developers or continue to study adaptivity unless we can present techniques that have been tried and tested through user evaluation studies. This research attempted to develop and combine complementary techniques that would achieve the following; reduce cognitive overload, provide dynamic navigational support, reduce information overload and not hide links or content for any level of user. This led to the development of the shading method of adaptive presentation.

The first empirical investigation considered the static sections of the SaDy SM and some dynamic variables. This was done by comparing three applications: a shaded text application, an adaptive navigational support application and a pre-existing hypermedia application that was authored and developed by archaeologists. The first study revealed some important variables of the SaDy SM and also revealed the usability of the shaded method of adaptive hypermedia. This study was then followed by a second empirical investigation, in which the shaded method of AH was combined with dynamic link annotation. This application was compared to a dynamic link annotation application and the existing hypermedia application non-adapted version of the application. Again, this study suggested the usability and acceptance of the shaded method of AH combined with dynamic link annotation and the importance of some variables in the SaDy SM.

1.2 Declaration

The work in this thesis is all the author's own work except the material used for the experiments which was provided by the Department of Archaeology at the University of Southampton and Dr. Ian Heath wrote the Microcosm filter that was needed for the second AH experiment.

1.3 Structure of the Thesis

The overall aim of this thesis is to present research which has been carried out in the area of AH to date and discuss future directions. The definition of 'adaptive hypermedia' can vary considerably. For example, what does adaptive mean, what can be adapted etc. To clarify this, Chapter 2 discusses adaptation in depth and defines what adaptation within this research means. Chapter 3 takes a look at evaluation in general and then presents some evaluation examples of existing AH systems. Chapter 4 takes a deeper look into the role of user modelling in AH. This chapter then goes on to present a survey which was carried out to construct a student model framework for educational adaptive applications. The first adapted hypermedia experiment is discussed in Chapter 5 and the results are examined in Chapter 6. Chapter 7 discusses the development of a second empirical study, which was carried out to investigate AH techniques. These results are discussed at length in Chapter 8. Chapter 9 presents the conclusions from the study and provides some possible directions for future research.

Chapter 2 Introduction to Adaptive Hypermedia

2.1 Introduction

The main aim of this chapter is to present the concepts and techniques in Adaptive Hypermedia (AH). Section 2.2 begins by discussing AH concepts it then goes on to describe AH techniques which are classed as either link level adaptation or content level adaptation. Section 2.3 presents some examples of AH systems which support these various techniques. Section 2.4 introduces open hypermedia systems and discusses Microcosm as an AH system.

Before attempting to discuss AH and open hypermedia systems, it is important to clarify the definition of hypermedia again. Basically hypermedia allows us to navigate through an information space using associative links. Hypermedia provides us with multiple paths through an information space whereas traditional media such as books present a linear path through the information.

Within the scope of this chapter we are concentrating more on AH, however a detailed discussion on hypermedia systems and applications can be found in Lowe et al.'s book (1999).

2.2 Adaptive Hypermedia Concepts and Techniques

Adaptive systems recognise that individual users behave differently when using hypermedia systems. For example, experts generally know more and manipulate their knowledge in a different way, they want more choice, less guidance and they have better organised memory in general, whereas a novice requires more instruction and guidance and a more tolerant dialogue with inflexible structures (Nickerson, 1986) (Nicol, 1990). This illustrates how each user group has different needs and preferences. Adaptation has the ability to target different dialogue for different users and prevent particular user groups from using certain commands and nodes. AH has the ability to accommodate the variability between users as well as the changing requirements of the user over time, thereby presenting an application that is better matched to user needs.

Again, the most appropriate introduction to AH was found from this basic quote, which has been taken from Brusilovsky (1998b).

AH systems apply different forms of user models to adapt the content and links of hypermedia pages to the user.

During the literature review it was found that within the context of AH, authors refer to adaptivity at the user interface or application level. Schlingbaum (1997) refers to the various forms of adaptation at the user interface level, whereas Beaumont et al. (1995) refer to whole hypermedia systems/ applications. It is believed that these two areas of adaptation should be viewed separately as they have their own methods and ways of adapting, both as important as each other. Adaptation in user interfaces aims to provide usable information for overall access to the system whereas adaptation in hypermedia pages aims to provide useful information by providing content and navigational support (Nill, 1997) (Fink et al., 1998). However, it should be made clear that users interact with the interface as well as the application directly via buttons/ links (the links are a form of interaction therefore can be viewed as part of the interface). Therefore the interface and the application (i.e. the structure of the links and the domain model) should work together to aid and assist the user to achieve their goal. However, within the scope of this work we concentrate on the adaptation at the application level i.e. link and content level.

The various techniques in AH fall neatly under one of two main headings, content level adaptation and link level adaptation i.e. adaptive presentation or adaptive navigational support (ANS). Adaptive presentation can be used to support groups of users with varying knowledge levels and adaptive navigational support is used to prevent users from getting lost by annotating links and providing guidance during navigation (Brusilovsky, 1998b). Examples of AH systems using these techniques are described in section 2.3.

2.2.1 Various Dimensions of Adaptation in Time

Adaptation can be viewed from 3 dimensions in time i.e. at points in time when the adaptation takes place.

1. Adapted - where the hypermedia system is adapted for different groups of end users at design time e.g. novices, regular and expert users.
2. Adaptable - where the end user has the ability to change the functionality and or characteristics of the hypermedia system.

3. Adaptive - where the hypermedia system adapts dynamically during the session according to end user behaviour.

For AH systems a stereotypical user model should be used during design time, where application independent characteristics can include; user preferences, capabilities, motor skills and application dependent characteristics including goals, knowledge of system and application. This stereotypical user model must be acquired as a result of comprehensive requirement analysis (Schlingbaum, 1997). However, if the hypermedia system is adaptive then the user model is built during interaction and dynamically updated. User models are discussed in more depth in Chapter 4.

2.2.2 Why Adapt Hypermedia?

In general, all hypermedia systems can be considered as adaptable to some extent i.e. they give the user the ability to make individual decisions. End users can adapt the system to their individual needs by deciding which links to follow next, for example clicking on textual links and bypassing graphics or other media. However, in general, most hypermedia systems provide the same sets of links and documents to all types of users, whereas in reality, user goals, preferences, background, knowledge and objectives differ when using the same hypermedia system. AH attempts to solve this problem for different class of users.

- AH can adapt information and links by using information about the user
- AH can reduce the possibilities of getting lost and confused, thereby reducing the 'lost in hyperspace' phenomenon by providing ANS and limiting the browsing space
- AH can provide adapted comments and help to individuals

Using AH resources within the educational domain results in other benefits including:

- Overcoming of resource problems e.g. lack of staff (Underwood et al., 1996)
- Aiding of special learners
- Presenting information in various ways
- Easing load on teaching staff
- Reducing the number of teaching hours (Orsini Jones et al., 1996)
- Allowing self paced learning by providing an environment for one-to-one correspondence between student and application tutor

Evidence also suggests that adaptive systems can increase the effectiveness of a user interface (Cakmakov et al., 1993) and that interface characteristics can heighten or reduce cognitive overheads (Wright, 1991). However, it can be argued that adding intelligence and control to an interface is contrary to the philosophy of hypermedia; which is supposed to give the user full control to explore the available content.

2.2.3 Adaptive Presentation

Adaptive presentation in the main refers to changing the presentation of the content of a document or style of the text depending on a user model. For example presenting different groups of users such as expert and novice groups with a different version of the same content. Novice users may get a brief introduction or more explanations and experts may be provided with more detail or complex information. At a very basic level we can say that adaptive presentation is based on 'fragment variants' which may be chunks of text or whole pages.

An example of an adaptive presentation methodology is conditional text, which is used in ITEM/IP (Brusilovsky, 1992). In this method all the information is divided into chunks where each chunk is associated with a level of user knowledge. Each chunk is only presented if a user meets a particular requirement. Stretch-text is a method where chunks of information may be displayed or not, depending on the user's knowledge level. A traversed link results in an expanded explanation of the link, which replaces the hot word resulting in a longer version of the original document. This text can also be collapsed back to the hot word. MetaDoc (Boyle et al., 1994) applies this method. In this method the user or the system can make adaptation decisions, resulting in an adaptive and adaptable hypermedia technique. The most powerful technique for conditional test is a frame-based approach, used in Hyperadapter (Hohl, 1996). Where information of a concept is represented in a frame and slots of a frame can contain several different explanations of a concept. A technique developed and used in the PUSH system (Hook et al., 1996) is a cross between frame based and stretch-text.

2.2.4 Adaptive Navigation Support

Adaptive Navigation Support (ANS) refers to changing the presentation of links in a document, many methods exist under this heading, for example direct guidance, sorting, hiding and annotation.

A very basic method for ANS is to provide a different set of links for different groups of users. Direct guidance is also classed as a basic method for ANS, the system decides which link the user should follow next depending on their goals and preferences (as used in ISIS-Tutor (Brusilovsky et al., 1994)). Although this is a useful technique, it does require additional link management support (Brusilovsky, 1998b). There are two main ways to support this method, the links can be visually outlined or links can be dynamically added or deleted/ hidden. Link ordering/ sorting occurs when available links are re-ordered dynamically on the document usually with the most relevant first. This is a useful technique for information retrieval systems such as HYPERFLEX (Kaplan et al., 1993).

Link annotation augments links to provide the user with some idea of the status of the node behind each link within the current document. Annotations can be in the form of text or visual clues. Visual clues can be in the form of different icons, colours, font types and sizes. This method eliminates incorrect mental maps (i.e. how a user models the hypertext network in their mind) and is more effective than link hiding as hiding can only distinguish 2 states of a link; relevant or not relevant. Hiding links is the most commonly used method to support ANS. It works by restricting the number or presentation of links, which are presented in a document and can help to reduce cognitive overload. But it does result in users building incorrect mental models of the hypertext. De Bra et al. (1998a) present 3 variations of hiding, link removal which removes the non-relevant links, link disabling where the link is visible but the functionality is removed and link hiding where the functionality is retained but the link text is presented in the same font/ colour as the rest of the text (usually black). In the latter, the link can be used if the user discovers it. Dimming can be used to support link disabling, i.e. links can be dimmed by changing their colour to grey. One or more of these methods of adaptation can be used within one AH system. Examples of AH systems using either adaptive presentation and/ or ANS are discussed in the next section.

2.3 Examples of Adaptive Hypermedia Systems

2.3.1 AHA

In a paper by De Bra et al. (1998a) the authors present the AHA system (Adaptive Hypermedia Architecture) which was developed at Eindhoven University. AHA is an open adaptive hypermedia architecture for the Web, the difference is that it separates the interface from the user modelling and adaptive features. The system supports an educational hypermedia course on the Web, however it can also be used to support other hypermedia application areas such as on-line help, on-line information systems, institutional hypermedia and personalised views. AHA is one of the most versatile AH systems. It can support 4 types of link level annotation and 2 types of content level adaptation. The text is fragmented and the links are annotated by changing the colour of the link. This system is also adaptable in that the user can choose between link annotation and link hiding. Different fragments of text can be displayed or hidden depending on the user model. To support fragment variants in AHA, a preprocessor filters the content by conditions encoded into HTML comments as HTML cannot support fragments of text which may be visible or hidden. The paper provides examples of correct and incorrect code, which is useful to anyone with knowledge of HTML.

The paper discusses three sub groups of link hiding: link removal, pure link hiding and link disabling. However, AHA only supports link annotation and pure link hiding although link removal can be simulated i.e. the author or user can chose between link hiding and link annotation by changing the colour scheme. Link hiding is supported by selecting black as the colour for undesirable links, assuming the rest of the text is also presented in black. The user will only know the link exists if they run the mouse over the text at which point the cursor will alert the user to the existence of a link. Link annotation is provided by classing the links and displaying them in a different colour. The links may be desirable or not, this is where the pre-processor classes them as good or bad respectively with respect to the user model. It is possible to support adaptive link removal by changing the HTML code i.e. the anchor tag into conditional text. Direct guidance can be supported via conditional content however it is a more difficult process. For more detail of the AHA architecture refer to another paper by De Bra et al.

(1998b), which gives a detailed description of the AHA architecture and gives a very brief but solid introduction to AH and it's techniques. This paper explains the system in more detail and provides 3 examples of applications based on AHA. AHA is used to present an on-line course 'Hypermedia Structures and Systems'. Again the paper provides more detail by providing examples of the content adaptation and how the conditional text may result in hiding it or replacing it with more relevant version of the text.

This paper also goes into detail about the structure of each node as well as how an author can develop a user model. AHA can support up to 148 different types of users which is an impressive level of granularity. The authors suggest a simple but effective user model, i.e. (c, v) where c is a concept and v is a value. In this user model, the references are represented by Boolean values, for example true or false, which shows how much a user knows about that concept, i.e. known or not known respectively, but the colours of the links are stored as explicit colour values. The system also builds a log file for each user, where the time a user spends on a document is stored as well as each test score. In this system, knowledge is generated by reading pages and taking tests. The authors explain that the user model is limited in that if an undesirable node is read the model assumes that the user did not learn anything and the user model is not updated. However, it is explained that INTERBOOK (see section 2.3.2) has overcome this drawback by allowing each node to have four states. The paper also discusses the 3 link types that AHA supports; conditional, unconditional and external. An external link is one which points to a web page outside the current hypertext. A conditional link may be classed as good, bad or neutral depending on the user model. An unconditional link can be classed as neutral or good depending on if it has been visited or not. Again this paper goes into detail about the conditional content with example of code. The paper also presents a clear and simple diagram of the AHA architecture. The AHA engine is a CGI script and the paper sets out all the functions it performs, and how it works. AHA is classed as an open architecture because it has the ability for other systems to pick up user models as well as to update them. This is done by the system generating an external representation of the user model. However, the authors point out that this is only the first step on the way to building open AH systems. The paper concludes by outlining extensions to the system.

Another paper worth mentioning is De Bra (1999) which describes AHAM (Adaptive Hypermedia Application Model). This is a reference model for AH applications. AHAM is based on the Dexter model. AHAM augments Dexter with features for doing adaptation based on a user model, which persists beyond a session. The main aim of this paper is to define and place AH within the Dexter Model and also develop a framework to aid the development of AH systems in which authoring would be easier than in a system such as AHA and INTERBOOK. A diagram is presented to show how AHAM fits into and extends the Dexter Model. The paper provides the reader with 10 detailed definitions of the various components. For further information please refer to De Bra et al. (1999).

2.3.2 INTERBOOK

INTERBOOK (Brusilovsky et al. 1998a) (Eklund et al., 1998c) is an authoring and adaptive textbook delivery system based on the WWW. INTERBOOK provides dynamic ANS by using three different fonts and four different coloured bullets to represent four states; ready-to-be-learned, visited, not-ready-to-be-learned and unknown. For example the system either places a different coloured ball next to a link to show it's current status, or one of three levels of check marks to represent already visited concepts. The adaptation takes place with respect to an individual user model, which is constantly updated. Eklund et al. (1998b) present a more detailed discussion of the system and presents a strong case for the use of AH in higher education.

2.3.3 ISIS-Tutor

ISIS- Tutor is one of the earliest AH systems (Brusilovsky et al., 1994) in which ANS support is provided by link hiding and annotation. It uses different colours and special marks to annotate the set of links leading from the current node to related nodes according to the users current knowledge and educational goals. This system holds a basic student model. Each concept and related hypernode is represented as one of four states; not-ready-to-be-learned, ready-to-be-learned, in-work and learned.

2.3.4 ANATOM-TUTOR

A paper by Beaumont et al. (1995) describes two AH systems within an educational context: ANATOM-TUTOR, which is an intelligent tutoring system, and ISIS-TUTOR.

In ANATOM-TUTOR, adaptation takes place at the link and content level. Adaptation at the content level involves the adaptation of the style and content of the text. This system also provides custom tailored explanations. The system utilises a user model, which uses stereotypes and mechanisms for deducing information about a user. User information is gathered using a questionnaire, which the user completes at the start of the first session. This is used to apply stereotype knowledge from one of two main groups of users, distinguished as beginner and advanced.

2.3.5 HYPERADAPTOR

HYPERADAPTOR (Hohl et al., 1996) is an adaptive and adaptable hypertext system which supports the learning of programming. The paper goes into detail about the architecture and explains the knowledge base in detail. It also gives a detailed diagram of the HYPERADAPTOR architecture. The system builds a detailed, dynamic student model of user expertise, which is then used to provide personal assistance. The model holds information such as personal data, programming competence and personal interests. The system supports four user stereotypes: novice, beginner, intermediate and expert. Link attributes are evaluated by selection rules, which prevent users being presented by information they are not ready for etc. The main features of the system are that it can adapt information and links for individual users.

Original user information is collected via questionnaires, which are filled out at the start of the first session. The paper provides screen shots of the questionnaire and goes on to present two screen shots of the content, one for a beginner and one for an expert. Topic browsers can be used to aid navigation. The system aims to incrementally broadening the user's knowledge by tracing their progress. The system is one of only a few which provide users with access to their profile, called a user model inspector, again a screen shot is provided. The last section of the paper discusses related research and talks about other adaptive systems in terms of their user modelling components. In conclusion the authors suggest extensions to the system by designing more realistic user models as well as taking into account the concept of decaying knowledge etc.

2.3.6 ELM-ART 11

Weber et al. (1997) developed a system called ELM-ART (ELM Adaptive Remote Tutor) which is an intelligent interactive textbook based on the WWW and supports the learning of LISP. It provides dynamic adaptation by annotating and sorting of links and adapting the presentation of the content. Adaptation decisions are based on a user model, which is updated dynamically during the session. All the documents presented are generated on the fly. The sorted links are presented with the most relevant first. ELM-ART 11 uses the traffic light metaphor to visually annotate the links. The green ball means that the link is recommended, a red ball means that this page is not ready to be learnt, the yellow and orange balls mean different things at different times. For example, the yellow ball can mean a question has been answered correctly or a particular document has been visited.

2.3.7 PUSH

Hook (1997) developed the PUSH system, which is a web-based system developed to aid users when using an on-line manual. PUSH does not affect how users navigate around the system only how much information is displayed on a document. PUSH uses stretch-text technology and selects what to hide in a page thereby avoiding information overload. Stretch-text is a method of self-adaptation where a highlighted keyword is a link, if clicked the system replaces the original keyword with an expanded explanation, within the same document. This text can also be contracted back to the keyword. Local maps of the information space are also available which can be used as a navigational aid.

2.3.8 AVANTI

Fink *et al.* (1997) designed and evaluated a system called AVANTI, which can be accessed via the WWW. It is a distributed information system that provides information about a geographical location to a variety of users, such as tourists, the elderly, the blind etc. The adaptability is provided at the user interface and the application level i.e. the content of the documents and presentation of the links. The user model holds information on user interests and preferences, domain knowledge, computing experience and their interactions with the AVANTI system. The system decides what information to display depending on the user model, which is updated during the interaction process. The original assumptions for the user models were acquired by interviewing and dialogue interaction from which users were allocated stereotypes.

2.3.9 HYPERFLEX

HYPERFLEX is an on-line information retrieval system developed by Kaplan et al. (1993). The HYPERFLEX system implements adaptive sorting of links, where the system provides the user with navigational support by displaying an ordered list of nodes related to the current node, depending on the relevance of each node. The most relevant is displayed first.

2.3.10 GUIDE

The Guide system (Cheverst et al., 2000) is a live system that provides visitors with up to date information about a city using a hand held wireless unit. Information presented to visitors is tailored according to their personal context (location and profile) and the environmental context. The visitors profile is constructed at the start of the session where they enter some personal details such as interests and language. This information is currently static but work is being carried out to make it dynamic.

2.4 Open Hypermedia Systems

Earlier hypermedia systems were mainly classed as closed systems, i.e. a system which provides fixed applications which are highly integrated with the links mechanics of the system (Legget et al., 1993). The documents contain the embedded links which makes external access of documents and links by other hypermedia systems impossible. On the other hand, open hypermedia systems allow any application to manipulate the hypermedia mechanics and functionality of the system. However, the level of openness may vary depending on the restrictions of the protocol. An open hypermedia system should be open with regard to size, the ability to import new nodes, links and anchors without limitation. The system should also be able to import and use any data format and allow access to the links service by any application. In addition, the hypermedia data model should be configurable and extensible, it should be possible to implement the system on various distributed platforms and finally the system must support multiple users and allow them to hold individual views within the system (Lowe et al., 1999). With these criteria, no system can claim to be truly open and the WWW must be classed as closed, however it is possible to implement an open hypermedia system within this type of environment for example the Distributed Link Service (Carr et al., 1995). A good introduction to open hypermedia systems can be found in the introduction to the JoDI (Journal of Digital Information) special issue on the subject (Wiil, 1998).

2.4.1 Microcosm as an Adaptive Hypermedia System

Microcosm is an example of an open hypermedia system. It separates the links from the content and no mark-up is required. The main aim of developing Microcosm was to reduce the effort required for authoring. Firstly authoring large sets of documents is a difficult task anyway and secondly, embedded links once created are difficult to keep track of. Also these links are only position dependent and cannot be reused. Another important aim of the project was to enable the system to apply links to read-only media. The resulting system, does not mark up the content or data and uses a link service, which can be applied to any data type and data within a third party application. The separation of the data and link structure allows the reuse of content. Again this is important for AH systems, if links are not embedded, the system can reuse the same content to display to different users in various combinations and contexts.

The Microcosm architecture (see Figure 2-1) is based on the use of filters, these are programs or processes that can communicate with Microcosm, and handles messages with the Filter Manager. Almost all process are designed and implemented as filters (Lowe et al., 1999). This allows unlimited scope for adaptability, and filters can be designed to carry out AH functionality and support various AH techniques. One such filter may be used to support ANS by displaying annotated dynamic links at run time based on the content or context of a source anchor rather than the location within a document. Another example is the computed links filter, which applies information retrieval techniques on a selection of text on other text documents to find the same or similar vocabulary.

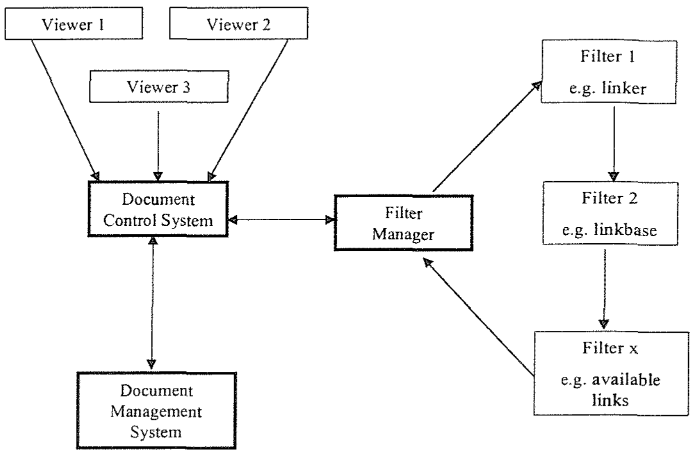


Figure 2-1 The Microcosm Architecture

All links in Microcosm are stored in a linkbase. Multiple linkbases can be used at any one time allowing for different views of the same content and potentially resulting in an adaptive application. By separating the links into separate linkbases the author or user can tailor the information space for different groups of users. Again, these links may be configured to support ANS by displaying a different colour for different types of links. This also allows Microcosm to support the use of hidden links i.e. by displaying the links in black to match the document text. Further annotation techniques such as adding icons etc can also be supported by designing new filters.

Linkbases support three link types:

specific link -is a link from a particular object at a specific point in a source document that connects to a particular object in a destination document

local link -is a link from a particular object at any point in a specific document that connects to a particular object in a destination document

generic link -is a link from a particular object at any position in any document that connects to a particular object in a destination document

It is recommended that a minimum of three linkbases are used for any application, one holding links created by the author, the second holding generic links to background material and glossary terms and the third holding links which have been created by individual users. This environment is ideal for developing AH systems and applications, as all information requires a basic structure connecting the background concepts, while generic links, can lead to explanations presented in different ways depending on the end user. Also the context dependent links can be generated for different types of users or contexts depending on the end user requirements and goals.

Microcosm also allows users to become authors by giving them the ability to create links within existing hypermedia documents, without changing the basic structure provided by the original author. This adds a new dimension to AH systems where users can literally create their own personalised views of the data depending on their personal goals and knowledge levels. This can also be viewed as adaptable hypermedia, where the user has the ability to make the system adapt to their personal mental maps.

Three primary navigation or browsing methods exist within Microcosm, the hypermedia link service, computed linker and the Document Management Service (DMS) which is a file browser. Again this can be viewed as adaptable as the user has the ability to decide which navigation method to use or alternatively the system can dynamically decide which method to present to the user depending on their characteristics. One aspect of AH systems which Microcosm is currently lacking is the provision of a user model engine, however, this can be overcome by designing and implementing a user model filter.

This section has briefly discussed how Microcosm can be viewed as an AH system. A more detailed discussion of the Microcosm architecture can be found in Davis et al. 1992. The following quote (Lowe et al., 1999) summarises Microcosm within the AH system paradigm.

> The Microcosm architecture enables different views of the same set of documents to be presented to different users either by associating different logical types to the files in the application in the DMS or by using different combinations of linkbases. The filter architecture allows for the dynamic generation of links via an algorithm defined in a filter, which can in turn lead to the development of AH applications.

Having discussed the various types of AH systems, in the next chapter we consider how their use can be evaluated.

Chapter 3 Evaluating Adaptive Hypermedia

3.1 Introduction

Evaluation is an extremely important part of any project lifecycle, especially where end-users are concerned. Many different types of evaluation techniques exist, each incurring varying monetary, human and time costs. Results from evaluations can be used to assess a given design solution, compare designs or as a foundation to base future design decisions. A definition given by Jacobs (1998) summarises evaluation;

> Evaluation is the systematic reviewing of the value and effectiveness of a piece of courseware, and is normally carried out by collecting experimental and other data before coming to objective conclusions.

To investigate the evaluation of AH it is necessary to explore the various techniques used in the evaluation of user interfaces in general, as well as issues involved in evaluating hypermedia. This chapter presents a three-part review of evaluation techniques, starting with a brief look at usability issues in section 3.2. Section 3.3 discusses general techniques used in evaluating user interfaces. Section 3.4 discusses the evaluation techniques used in evaluating hypermedia applications, and section 3.5 discusses the evaluation techniques used when evaluating AH applications.

3.2 Usability

Usability refers to the effectiveness of the interaction between people and the system, i.e. all the activities performed by the user that are affected by the technology they use. Individual aspects of usability are called usability factors, for example, speed of performance, incidence of errors, recovery from errors, retention of learned skills and satisfaction (Newman et al., 1995). Most usability measures depend on the activity being performed. Usability can be measured analytically by performing simulations of user activities. These techniques can be carried out quickly and informally and they do not involve real users. Usability can also be measured empirically, by observing and recording actual user performance of frequent or critical tasks usually performed on a prototype. These techniques require prior planning and careful experimentation.

3.3 Interface Evaluation Techniques

3.3.1 Cognitive Walkthrough

The cognitive walkthrough method involves individuals or groups of evaluators inspecting a user interface by completing a set of tasks and evaluating its understandability and ease of learning. The user interface is often presented in a paper mock-up or working prototype. This method of evaluation works well with team walkthrough situations compared with individual walkthroughs (Karat et al., 1992). This technique is usually applied to designs for systems that will be used by people without prior training or for exploratory type learning environments. It can also prove to be most useful in systems that have been changed or extended (Polson et al., 1992).

A study by Wharton et al. (1992) found that the walkthrough method required a higher knowledge of cognitive science terms, concepts and skills than most software developers would possess. Other problems with this method are that only a small subset of the tasks can realistically be evaluated, task selection itself may be problematic and the evaluator does not have a global view of the interface. However, the walkthrough method is good if resources are limited and may be useful when used early in the development lifecycle, as it could be used to evaluate competing solutions before development begins. However, it is felt that this method is not suitable for use on highly functional systems unless further extensions are made to the method (Karat et al., 1992) (Wharton et al., 1992).

3.3.2 Checklists

The advantages of using a checklist approach to evaluation are that it provides a structured list of relevant criteria for evaluation and can replace expensive empirical testing. This method is easy to handle and it can produce the impression of a complete set of evaluation criteria and procedures. However, it has been suggested that evaluation based exclusively on checklists can only be considered as a 'first level review' and would therefore also require support from a variety of other evaluation methods to produce a complete evaluation report.

3.3.3 Interviews

Interviews are a good technique when trying to capture in-depth information such as attitudes, impressions, opinions and ideas (Dix et al., 1993). The interviewer can direct responses or develop responses further if the they feel they do not understand, resulting in a more precise and controlled discussion. The face-to-face method can also be useful because the interviewer can pick up on physical responses or reactions of the interviewee.

3.3.4 Questionnaires

Questionnaires, in which users describe their experiences, attitudes and opinions of using the system, gather subjective data from users. Some evaluators argue that this is the most important type of information i.e. knowing what the user thinks is more revealing than objectively measuring the same parameter. Questionnaires can be designed to provide specific or general responses by using open ended or closed questions. They provide an accurate and fast method of gathering data. Subjects may feel more inclined to respond honestly to questionnaires as they are usually anonymous. Questionnaires are especially useful when requiring a large number of responses. Once created they can be handed out and collected at a later time and can be distributed by inexperienced people at a low cost. The data from the results provides a good basis for using statistical analysis. Many types of questionnaires can be used. For example pre and post-task questionnaires can be used to establish the acquisition of knowledge and pre and post-session questionnaires can be used to establish subjective attitudes about particular system issues. Lindgaard (1994) highlights five issues regarding the design of questionnaires; the format and content of questions, response categories and scales, coding of responses, layout and structure. A major drawback of using questionnaires is that they are less dynamic than interviews and responses can not be discussed.

3.3.5 Focus Groups

Focus groups are designed to collect information by questioning 6-9 users who have been brought together to discuss the system. The investigator prompts a discussion by asking a few questions. This encourages other group members to speak out and may trigger other thoughts or comments. Again, responses may be investigated further or the investigator may ask for clarification of points if they do not understand a response. However, it is easy to lose track and discuss issues that may not be relevant.

3.3.6 Heuristic Methods

Heuristic evaluation is carried out by specialists who interact with the interface and attempt to form an opinion on how good or bad it is. Heuristic evaluation is an informal method of detecting possible usability problems using a set of guidelines called heuristics. In general, this method is a two-stage process. First it employs a team of evaluators and then heuristics are designed which are used to guide the evaluators. Examples of simplified usability heuristics are; be consistent, provide feedback, use simple and natural dialogue, minimise user memory load and prevent errors (Nielsen et al., 1990). Heuristic evaluations (or usability inspections) usually only involve expert evaluators as they are better at performing heuristic evaluations than non-specialists (Nielsen, 1992) (Garzotto et al., 1997). These evaluators look at the design and investigate what is right or wrong about it. This technique can be carried out relatively quickly resulting in reduced costs and resolution of problems. It is intuitive to perform, no advance planning is required and it is suitable for use in the early stages of development (Neilsen et al., 1990). This method is the most useful when evaluating a running interface (Nielsen, 1993). However, several evaluations must be carried out for the maximum benefit and it is ideally carried out by experts. There is a danger that experts may miss potential end-user type problems and may focus on problems and not solutions. A heuristic study by Nielsen (1993) found that heuristic methods only find usability problems and do not address other phases of the usability engineering lifecycle. It was also found that performance estimates from heuristic estimation and GOMS analysis were highly variable. A study by Molich (1990) found that heuristic evaluation was difficult to carry out, but also found that by aggregating the evaluations of teams, the overall results did improve. Molich et al. stated limitations of the method implying that the approach only looked at individual usability problems in the development phase where the design is completed but needs refinement.

3.3.7 Empirical Evaluation

The most common empirical strategy is the usability test (user testing) where the aim is to study the performance of real end-users. Formative or summative evaluation methods may be used. Formative evaluation assists the process of enhancing the systems design by helping to form the solution to the design problem. Summative evaluation tests the design as a whole, so it is seen as being less effective than formative evaluation (Newman et al., 1995)

Summative evaluation is usually carried out on the final version of a system, with real end-users performing real life tasks in a real work environment. Formative evaluation is carried out with users to establish problems while it is still possible to make modifications.

Some researchers feel that experiments that take place in a laboratory can not apply to real learning environments. It is felt that by studying the user closely, this approach may lead to the Hawthorne effect, where the very act of performing the evaluation experiment affects the results; students may perform differently in an evaluative situation than in a normal learning environment (Jacobs, 1998). An empirical study carried out by Karat et al. (1992) identified the largest number of usability problems and a significant number of relatively severe problems that were missed by the walkthrough method. This finding was consistent with a study undertaken by Desurvire (1994).

3.3.8 Observational techniques

Observational techniques include direct observation, video recording and analysis, protocol analysis and system monitoring. Direct observation requires evaluators to watch the user as they perform their task, usually by sitting near the user and taking notes. It can be intrusive but this can be overcome by using a one way mirror. This technique should be used in conjunction with another technique. User observation is good at identifying external pressures e.g. social issues, which may impact the user during real time use. This is an expensive method, which requires at least one observer to observe each user individually, although it would work for a limited number of observations. A problem with this method is that it may be subjective unless there is a team of observers. A video camera may also be used to observe the user and gather information. Cameras may be hidden from the user to prevent the Hawthorne effect (Salzman et al., 1994). Video recording provides a full and permanent record of the session. The results can be correlated with other analysis data. The results can also be analysed by anyone at a later stage. A major problem with using video cameras is that analysing the data can be time consuming (Lea, 1998) (Nielsen, 1993).

Protocol analysis is a technique where the user is requested to think aloud, i.e. to provide a verbal account of their thoughts, feelings, opinions and actions when interacting with the systems (Jorgensen, 1990). Protocol analysis requires little training; therefore designers can perform the evaluation. It also provides qualitative data about the user's cognitive processes and data collection is fast. However, some people find this an unnatural action and therefore find it difficult to do (Kato, 1986). Analysing the results may also prove difficult and the results may not provide an accurate description of what is actually happening. In this case retrospective thinking aloud can be used once the evaluation session has finished. The advantage of this method is that it does not interfere with the evaluation session itself (Hix et al., 1993). However, this could lead to biased and selective recalling of events (Lea, 1988).

3.3.9 Software Logging

Performance data may be collected by a piece of software logging the user actions as they interact with the system. This results in data which is collected automatically and is reliable, therefore exact measures of performance can be extracted (Lean, 1988). Other advantages of this method are that the data is recorded unobtrusively, the data is highly accurate, the collection method is cheap and the evaluator does not need to be present. However, software logging can result in large amounts of data being generated where some systems may gather data at a very low level. Another major drawback is that the log files do not explain why users are performing a task in a particular way. For example time stamps for the time a user opened and closed a particular document may be registered at ten minute intervals but this does not imply the user was viewing the document for that total amount of time. Therefore this method must be accompanied by direct observation, audio or a video recording, to help assess when the user is actually looking at the screen, and when they are not.

3.4 Hypermedia Evaluation Techniques

Evaluation of user interfaces can be difficult enough. However, evaluating hypermedia or multimedia applications results in additional problems. This is due to the significant differences between the two technologies, i.e. the additional complexities of these technologies (Tergan, 1998) (Hutchings et al., 1994) (Nielsen, 1990).

Hypertext applications pose special problems because of the way users subjectively navigate the information base. Two other important issues are that the hypermedia based learning tasks usually involve more complex user-system interactions and that learners can actually construct their own knowledge while learning. This complexity of hypertext applications increases the importance of evaluation but also makes them less straightforward and therefore requiring different types of evaluation criteria (Wright, 1991).

Again, as with designing interfaces, iterative design using usability evaluation is an important aspect of the hypermedia lifecycle and is vital in producing a successful application (Hutchings et al., 1992) (Nielsen, 1990). However, there is a distinct lack of research in evaluating hypermedia applications and a greater lack of evaluations carried out with respect to educational courseware (Jacobs, 1998). It seems that more work has been carried out on the technology side of hypermedia systems (Nielsen, 1990).

The usability of a hypertext application is determined by a combination of the usability of the underlying hypertext system engine (the presentation and navigational support available), the usability of the contents and structure of the information base, and by how well these two elements fit together.

Nielsen (1990) suggests the use of usability testing via heuristic evaluation by modifying existing heuristics or creating new ones for hypermedia systems. He also suggests the use of questionnaires, interviews, logging and observation. Nielsen's (1990) discounted usability engineering approach presents five usability parameters- easy to learn, efficient to use, easy to remember, few errors and pleasant to use. Gillham et al. (1995) discussed Nielsen et al's (1990) original heuristics for usability evaluation, which suggest that the system should be simple and natural in language, speak the users language, minimise user memory load, be consistent, provide feedback, provide clearly marked exits, provide shortcuts, have good error messages and prevent errors.

A study by Nemetz et al. (1997) to compare evaluation methods by evaluating hypermedia found that heuristic evaluation discovered 54 problems and usability testing found 32. Both methods found a similar number of severe problems but heuristic evaluation found more simple problems. The authors argue that results show the superiority of heuristic evaluation over usability testing. They also report that only 20 problems were found in common by both techniques, implying that these are complementary techniques.

It has been suggested by Garzotto et al. (1997) that empirical and inspection evaluation methods are complimentary and should be used together. They present a model called Systematic Usability Evaluation (SUE), which attempts to reduce the disadvantages of both these methods. SUE attempts to adapt existing evaluation methods and introduce more hypermedia specific issues.

The cognitive walkthrough method has also been used to evaluate interactive systems (Newman et al., 1995). Johnson (1992) suggests that this is the 'easiest' method of evaluation. Analysis of a hypertext application using walkthrough involves simulating the way users explore and gain familiarity with interactive systems (Newman et al. 1995). An important issue when using the cognitive walkthrough method for hypermedia is that it must cover all the possible routes that the user may take when interacting with the system (Newman et al. 1995). However, this also makes the whole process extremely complex especially where the World Wide Web is concerned (Thimbleby, 1996).

Empirical evaluation in which a prototype is constructed and tested by users is seen as an essential part of interactive system design (Newman et al., 1995). Surveys, questionnaires etc. are widely used in evaluating hypermedia systems to gain subjective feedback from users of hypermedia systems (Hill et al., 1997) (Wills, 2000) (Makrakis et al., 1998) (Chanier, 1996) (Laurrilard, 1998). Other systems on which these methods have been used include, the Camille Project (1996) and the Perseus Project (Marchionini et al., 1994). User observation raises some important points during evaluation and highlights the importance of carrying out the evaluation on a fully operational system (Beaulieu et al., 1995) (Chanier, 1996). Again the Perseus Project and the Hypermedia Cammile Project used this method of evaluation. The checklist approach, although not as common as the other methods, has also been used as a means of evaluating educational hypermedia (Tergan, 1998).

Automatic session logging is a method which is used quite often in evaluating hypermedia applications (Nielsen, 1990) (Gillham, 1995) (Hutchings et al., 1994) (Hill et al., 1997). Even where it has not been used due to particular circumstances, this method is suggested by authors as a useful evaluation technique (Beaulieu, 1995) (Nielsen, 1990).

Even considering the complexities that characterise hypermedia, it is essential and possible to carry out a usability study (Neilsen, 1990). However, it is remarkable that the evaluation techniques used to evaluate hypermedia applications still rely on a combination of existing conventional methods, such as questionnaires, direct user observation, interviews and heuristics methods (with some refinement).

Although Desurvire (1994) suggests that different techniques may detect the same usability problems, Nielsen (1993) suggests that several well chosen complementary techniques should be used for evaluation (Hill et al., 1997). Petrelli (1999) suggests that it is not vital to use all these techniques and that it is sometimes more useful to use the cheapest.

In summary, we can say that most evaluations carried out on hypermedia systems use summative evaluation on systems that are fully implemented and operational (Jacobs, 1998). Existing evaluation studies carried out on hypermedia applications use at least two techniques with some studies using up to seven (Nielsen, 1992). The most commonly used techniques include questionnaires and user observation along with other methods of evaluation such as heuristic evaluation techniques.

3.5 Adaptive Hypermedia Evaluation Techniques

The previous sections discussed various methods used to evaluate interfaces and highlighted the differences in evaluating hypermedia, which arise due to the integration of the usability of the applications and system. However the situation becomes more complex when we start to look at evaluating adaptive applications, systems and techniques. Very few studies have been reported in this area and even less which use a formal evaluation approach. One of the most comprehensive reviews on empirical studies of AH to date can be found in Brusilovsky's (1998b) paper titled 'Efficient

Techniques for AH'. The paper gives an in-depth discussion on various AH techniques, and then presents some of the most significant empirical studies undertaken within the field of AH. The following sections of this chapter will present examples of evaluation studies carried out in the field of AH.

3.5.1 AHA

De Bra's (1999b) paper provides brief examples of other AH systems, which support the various types of techniques and the outcome of informal or formal evaluation studies. This paper does not provide a discussion on an evaluation study carried out on AHA but does state informally that user feedback showed that users do not like link removal as they cannot see what lies ahead. Although no formal evaluation study is reported it does state that AHA is a good system to carry out these usability studies due to the number of adaptive techniques it can support. The authors also mention two studies carried out one using ISIS-TUTOR (see section 3.4.3), which found this method an effective way to reduce the number of navigational steps, and another study by Kobsa et al. (1994), which showed that displaying disabled links was more effective than pure link hiding. The authors also touch upon feedback from a course run earlier where users preferred link removal over link disabling, although this was not particularly well accepted either.

3.5.2 INTERBOOK

Two papers by Brusilovsky et al. (1998a) and Eklund et al. (1998c) introduce INTERBOOK, and one of the most complex evaluation studies carried out in AH. The evaluation study involved 25 undergraduate teacher education students in computing. The aim of the study was to assess the impact, if any, of user model based link annotation on student learning and user paths. The hypothesis was that adaptive link annotation would provide students with more efficient paths through the hypertext with improved learning methods. During the evaluation sessions, tests of knowledge were carried out as well as audit trails and questionnaires. In the first session, the system was used by the subjects, who then answered questions on its features. In the second session, users were divided and exposed to the database section of the book, 12 students were exposed to the system with adaptive link annotation and 13 without adaptation. In the third session users took a multiple choice test on the database section of the book and were then exposed to the spreadsheet section of the book. This time the groups swapped, the ones who had been exposed to the adaptive system were given the system without adaptation and vice-versa. In the last session, users were given a multiple-choice test on the spreadsheet section of the book as well as a questionnaire. The results presented in this paper are very comprehensive. The paper also presents the questionnaires given to the subjects as well as the mean and SD (standard deviation) of the results. The test scores were used as a measure of users knowledge acquisition.

The first analysis showed that link annotation had a negative effect on the database session and not a great deal of difference was found for the spreadsheet session. A further investigation into the audit trails revealed that some users already had prior knowledge and most of the navigation was done using the 'continue' and 'back' buttons, these were non-adaptive. It was found that users had very different navigational behavior, some never used the annotated links and others used them frequently.

The users were then analysed by considering those who navigated using annotation. It was clear that some users were only selecting nodes which were not recommended. It was also the case that users accepted the adaptive advice at varying levels. Brusilovsky et al. suggest a measure of acceptance of advice, which found a high positive correlation between the agreement rate and the database test score.

$$\text{Agreement rate} = \frac{\text{Number of green balls} - \text{Number of. Red balls}}{\text{All hits.}}$$

The results were also analysed with respect to the navigational tools that were used. It was found that ANS encouraged novice users to navigate in a non-sequential method of navigation. It was also found that users without ANS used more of the sequential navigation features. By the same token, users who did have ANS did tend to use the adaptive advice. This showed that users did trust the advice provided and users are willing to follow non-linear paths even if they are not experts. The results were also analysed by looking into the number of hits on different document states. This analysis revealed that users preferred to visit 'ready-to-be-learned' pages. It was also found that users also spent twice as long on these pages compared to other pages.

In conclusion the authors state that ANS is a useful AH technique, especially to new users and for those who trust and use it. However, it does introduce cognitive overhead, which may lead to distraction from the content. Results from this experiment are interesting as it was found that the students who used the adaptive link annotation performed significantly worse. However, after further analysis it was found that the more the students followed the systems suggestions the better they performed. The paper concludes by stating that much more work needs to be done in this area as there is a lack of formal experimentation.

3.5.3 ISIS-Tutor

Brusilovsky et al. (1994) carried out their experiment using ISIS- Tutor. The goal of the study was to check the efficiency of combining link hiding and annotation by using special markers within the educational context. The student model was used to distinguish one of four knowledge states or zones, namely 'not-ready-to-be-learned', 'ready-to-be-learned', 'in work' and 'learned'. Highlighting these zones visually was expected to help the user in navigation. For this experiment, 26 subjects took part, who were from the first year computer science course, at Moscow State University. They were all given ten questions to answer in 45 minutes after a brief introduction. The subjects were divided randomly into 3 groups, Group A worked with hypermedia without adaptation, Group B worked with AH (visual annotation) and Group C worked with non-restrictive AH, which supported link hiding and visual annotation. User actions were recorded and then analyzed with respect to the following variables, time to complete the course and the overall number of navigation steps. It was expected that the time and the number of steps would be less for the AH applications.

The study showed that the number of navigation steps and the number of transitions from concept to concept were less for the AH application and even less for the non-restrictive adaptive hypermedia application. In conclusion both AH applications supported efficient adaptive techniques (the ANOVA, analysis of variance test was used to check the significance). These techniques increased user performance by reducing navigation problems. The authors suggest that adaptive presentation can reduce the time for learning and increase comprehension but not reduce the number of nodes visited.

3.5.4 ELEM-ART 11

Weber et al. (1997) carried out an empirical study to look at the effects of combining different types of ANS with the guidance offered by the 'next' button using ELM-ART 11, which supports the learning of LISP as described in section 2.3.6. The traffic light metaphor was used to annotate links and the text was presented in different styles to aid colour blind users. A user model stored the knowledge of the user. Two treatments, each supporting two levels of AH were investigated simultaneously. The first treatment consisted of ANS by simply annotating the visited links with a yellow ball and not visited links with an orange ball. The second type of annotation used in the control group was similar to WWW browsers, which annotate links, which have already been visited. The second treatment contrasted the provision of a 'next best step' button with a version without this button.

Each user was randomly assigned to one of the four conditions. A questionnaire was presented to them, asking about their knowledge on programming languages and LISP. This experiment was only aimed at users without LISP knowledge. The first hypothesis was that ANS and individual curriculum sequencing with the 'next' button would motivate users to proceed with learning. Results were taken from 33 subjects. The results showed that users who used the 'next' button and did not have previous knowledge of LISP worked on about 10 pages more than the subjects without the button. This indicated that individual guidance provided buy the system was especially helpful for complete beginners.

The second hypothesis stated that the number of navigation steps are reduced by ANS and individual sequencing with the 'next' button. Although only 14 subjects took part in this study it was found that adaptive link annotation did not have any effect on the number of navigation steps, it helped the users at the beginning but as they got comfortable they followed the best path without any guidance. These studies showed that ANS and individual sequencing using the 'next' button are useful techniques in the starting phases of a session and for beginners.

3.5 5 AST

A paper by Specht (1998) titled 'Empirical Evaluation of Adaptive Annotation in Hypermedia' discusses the design and results of two experiments which compared different forms of AH. The first study used a hypertext on prionic diseases, which supported adaptive annotation and incremental linking of hypertext. In the second study, a WWW based learning environment on statistics called AST (Adaptive Statistics Tutor) was used to compare different forms of adaptive annotation.

In the first study, learning with three forms of adaptive hypertext was compared to a static hypertext. Using a combination of two AH methods, namely adaptive annotation and incremental linking, resulted in four different applications. Application 1, supported incremental learning and adaptive annotation so users were restricted in their freedom to navigate but could see where they would be allowed to go in future. In Application 2, only links relevant to user's current knowledge were presented as links. In Application 3, all the links were presented from the start and so there was no restriction on the navigation and the links were annotated so users had an idea of their relevancy. Application 4 was a static hypertext with no adaptation. Visited concepts were represented by a hook, not ready to learn concepts were represented by a red ball and ready to learn concepts by a green ball.

In the study, the learners knowledge was used to adapt the presentation of coloured balls in front of each link. Acquisition of user knowledge was assumed simply by the system when the user visited a concept. 85 users took part in the study. Before the evaluation session users were required to complete a demographics questionnaire and a knowledge test. At the end of the session the system automatically presented a post session questionnaire including questions about the usability and helpfulness of the adaptive methods. The time to read the nodes and the number of correctly answered questions were recorded by the system. This study showed that the best method to improve the acquisition of knowledge is a combination of adaptive annotation and incremental linking.

In the second study, different forms of adaptive annotation were compared using AST. Three applications were compared, the Annotation Application, which supported

ANS using coloured balls to indicate information about a link, the Hide Application which was the same as the Annotation Application but the not ready links were hidden, and the Static Application in which annotations were presented using white balls and hooks, so users only received information about concepts which they had learnt (hook symbols) and needed to learn (white balls).

Again users had to complete a demographic questionnaire and a knowledge test. Over a period of three months 180 subjects worked on the AST system, however results from only 67 of the subjects were taken into account. This was due to the complexity of the learning environment. The results suggested that the type of adaptive annotation can have an affect on the learning process, style of learning, consequences on motivation and the acceptance of the learning environment.

3.5.6 MetaDoc

The most comprehensive study carried out on adaptive presentation was with a system called MetaDoc (Boyle et al., 1994). The goal of the study was to analyse the use of stretch text. To carry out the experiment three applications were created. One without adaptivity, one which supported stretch text and another which supported adaptive stretch text. The subjects were from the Computer Science Department and were divided into three groups. Each group was given eight reading comprehension tasks and five navigation based tasks.

ANOVA was one of the statistical tests used. A significant difference was found between the reading comprehension time for stretch text and MetaDoc. For reading comprehension correctness and search time, a significant difference was found between the non-adaptive and the other two applications. However for the navigation variables such as search correctness, number of nodes visited, number of operations etc no significant difference was found. Reading comprehension time also decreased while understanding increased. Also it was found that adaptive presentation did not affect navigation i.e. no difference was noted between any of the applications with respect to the number of nodes visited. In conclusion it was found that stretch text was an efficient adaptation technique, which increased user performance by improving reading comprehension.

3.5.7 PUSH

The paper by Hook (1997), 'Evaluating The Utility and Usability of an Adaptive Hypermedia System' discusses an evaluation of the PUSH hypermedia system by comparing it to the non-adaptive version of the same system. In the study the following issues were evaluated: how successful the subjects were in retrieving relevant information, whether users were influenced by the adaptive system, if adaptivity reduced the number of actions and lastly the user's subjective preferences for the adaptive/ non adaptive system.

The study was carried out in a laboratory where the subjects were videotaped and actions logged. Statistics were recorded such as task completion time, actions carried out and inefficient use of the system. Nine subjects were used who all had WWW and hypermedia experience. They were asked to carry out five tasks and were asked to answer questions on their background, current understanding of some concepts and their preferences on adaptive or non-adaptive systems. The analysis found that the amount of navigational action between pages was not significantly different in the adaptive and non-adaptive systems. This confirmed that the system only affected the within page actions not the navigation between nodes. It was also found that the users trusted the decisions made by the adaptive engine and that most users preferred the adaptive version, as it required fewer decisions from the user.

3.5.8 HYPERFLEX

Kaplan et al. (1993) carried out a study, which looked into the efficiency of sorting as an adaptive technique using the HYPERFLEX system. The first study was designed to test the usefulness of goal directed searches. Only four subjects took part in the study, they were given 10 questions; five with system supported goals and five without. The results showed that 'goal based' adaptive sorting decreased search time and the number of searched topics. The goal of the second study was to compare the efficiency of two adaptive techniques; 'current node' based adaptation and 'current goal' based adaptation. Three applications were developed; one which only supported 'goal based' adaptation, one which only supported 'node based' adaptation and one which supported both. 18 subjects took part in the study. Six subjects were allocated to each application and each were given 4 search tasks. This study measured the time to complete each task and the number of nodes visited for each task. The results showed that the users of the application which supported both types of adaptation performed better than the other two groups. In conclusion although the number of subjects was small, the study did reveal that sorting can improve performance in search tasks.

3.5.9 AVANTI

Fink *et al.* (1997) evaluated the AVANTI system (see section 2.3.8). The system was evaluated at three different sites in three different scenarios. The aim was to establish if the system was beneficial to users and if it was technically feasible. The first experiment used 60 citizens and tourists including people with motor disabilities. The tasks given were to plan trips to various places within the area. The next experiment involved 11 tourists and then 40 tourists and business travelers. Their tasks were to book a holiday in a cottage. The final experiment used 20 visually impaired users and 40 without visual impairments. They were citizens and tourists and their task was to find information on a particular site. Information was collected via user observation, interviews, questionnaires and log files. The most relevant findings were that users who reused the AVANTI system tended to understand the system and adaptation better. The adaptive features were used more frequently by computer, AVANTI and WWW experts and the motion impaired users found the system more useful than the able-bodied people. In conclusion it was found that the adaptive features were well understood, used and appreciated.

3.5.10 Help Component

Encarnacao (1997) carried out an empirical evaluation of an AH help component. Users with different backgrounds first filled out questionnaires on their assessment of expertise on the system and application domains. Users were then asked to carry out pre-defined tasks with and without the adaptive help component. Their dialogue with the application was then evaluated with respect to their expertise assessment. These results were used to develop the user modelling aspect of the system as well as estimating the usability of the application with respect to different user groups.

3.5.11 GUIDE

The GUIDE System (Cheverst et al., 2000) was evaluated in two phases, using an expert walkthrough and a field trial, with the aim being to establish the quality of the user's experience. Visitors were allowed to use the system as they wished and were not set pre-determined tasks. A time log was also taken and direct observation was used with users encouraged to think aloud. An interview was conducted at the end of the session to gather subjective opinions. Over four weeks, 60 people used the system that varied widely in their demographics. The results were very positive with most users liking the provision of the context aware navigation and information retrieval mechanisms. They also enjoyed using GUIDE and trusted the information provided by the system. Another important outcome from the study was that designers must be extremely careful in what information they regard as relevant and display given a particular context.

3.5.12 Multimedia Course

Another study carried out in this field worth mentioning is the one carried out by Da Silva (1998), although this study was not formal, it was a preliminary experiment that presented some interesting results. The study focused on adaptive navigation by using link hiding and a user model, which held student knowledge about each concept. A major drawback of the paper is that is does not state how the evaluation process was carried out however informal it may have been. However, it did present positive results with respect to what students did and did not prefer about the adaptivity in question. It was found that link hiding confused users and that visual annotation should be considered instead. It was also found that it is important to present users with the size of the application by the use of concept maps. Another very important point to emerge was that the difficulty level of the documents was not enough to measure user expertise about a concept and that another variable such as coverage should be used.

3.5.13 Issues Raised by the AH Evaluation Studies

Due to the recent development of the AH field there is a lack of documented experimental research within the area. Very few AH techniques have been empirically studied, some authors report formal studies whereas some provide informal feedback. Individual empirical studies have been carried out on various techniques, mostly ANS but with less emphasis on adaptive presentation. However, there is no report of multiple studies of the same technique (except hiding) and therefore we can only rely on the results of one-off studies. Most of the reported studies have been carried out on WWW based systems and very few within the educational environment.

All the studies used conventional evaluation methods even though it is realised that evaluating AH is quite different to evaluating interfaces and traditional hypermedia. When evaluating AH all of the following evaluation techniques used questionnaires, user observation and logging, as well as and interviews and heuristic methods. Although there was an importance placed on the evaluation of a fully operational system during a period of time, this may not always be possible especially when working on a research project and restricted by time. The literature review also highlighted that most studies of adaptive systems were comparisons of the system with and without adaptivity. An important issue to come out of the literature review was what to measure in AH systems, studies carried out by authors such as Boyle et al. (Hook, 1995), Brusilovsky et al. used task completion time as their evaluation criteria, which can be useful for some systems. Other criteria which have been used to measure evaluation include reading comprehension time (Boyle et al., 1994) in MetaDoc, number of nodes visited, number of times user revisited concepts studying, task completion time (Brusilovsky et al., 1994) in ISIS-Tutor and search time (Kaplan et al., 1993) in HYPERFLEX.

3.6 Evaluation Review Conclusion

It was interesting to find that each literature review brought out many similarities between what methods had been used for the evaluations. It was also interesting to find that not much in the way of methods, number of evaluators or motivation had changed over time. However, it was disappointing to find papers outlining the same problems at present as other authors had mentioned many years previously.

All the papers agreed in conclusion that several evaluators should be used to carry out the evaluations and that different types of evaluators were required to provide different outcomes. The authors also agreed that most of the evaluation techniques are complementary to each other, yield different results and recognise different usability problems. Therefore, dependent on the availability of resources, a combination of evaluation techniques should be used in an empirical study. The main differences found between evaluating interfaces and hypermedia was that hypermedia is more complicated due to the integration of the system with the application and hypertext functionality. The main differences found between evaluating traditional hypermedia and AH were the measures taken to establish usability, although the same evaluation techniques were used in all three evaluation environments i.e. interfaces, traditional hypermedia and AH.

3.7 Next Stage

Chapter 2 described various AH systems and the AH techniques they have implemented. This chapter discussed the evaluation studies carried out to evaluate interfaces, hypermedia applications and AH applications. The next chapter is presented in two main sections, the first concentrates on introducing and describing the concept of user models with respect to adaptive and intelligent tutoring systems. The second describes a study which was carried out to investigate the development of a basic framework for a user model to be used within an educational AH system. This study was carried out to develop a deeper understanding of AH systems and the importance of user models as a component of them.

Chapter 4 The Role of User Modelling in Adaptive Hypermedia

4.1 Introduction

AH systems, by definition are designed to recognise that different users behave and manipulate their knowledge in different ways. For example, novice users would require detailed instructions and guidance and would possibly be more tolerant of slower systems. AH systems and intelligent systems model the differences between users by developing user models (UM) or in the case of educational systems, student models (SM). This chapter is divided into two main sections. The first presents an introduction to user modelling and the second presents a survey, which was undertaken to investigate the construction of a basic SM framework for use in an educational AH system.

4.2 A Generic Adaptive Hypermedia System Model

Chapter 2 discussed various AH systems and adaptation methods they support. This Chapter will aim to present a generic AH model. At an abstract level the architecture of most educational AH systems must consist of at least the first three of the following components (Wu et al., 2000):

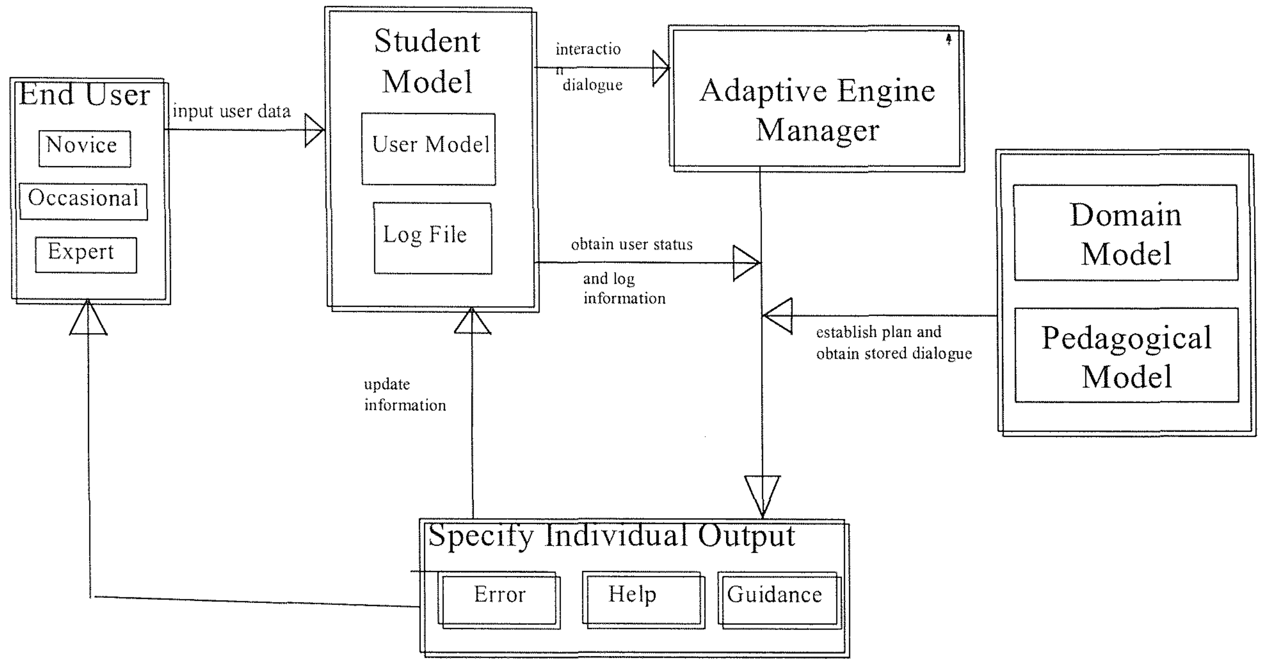
- Adaptive Engine
- Student Model
- Domain Model
- Pedagogical Model.

Figure 4-1, presents a generic model for a dynamic AH system and shows the relationships between the various components. The adaptive engine establishes the user's goals and is responsible for making adaptation decisions. For adaptive interfaces, the adaptive engine may consist of a dialogue manager as in Cockton et al.'s model (1986) and or a plan recogniser as in Toterdell et al.'s model (1987). The Adapts (Brusilovsky et al., 1999) architecture holds a separate adaptive diagnostic component which selects the most relevant task for the user. This is then passed to the adaptive engine, which provides the adaptive support.

The SM holds information on the user's knowledge of the domain and their progress. The SM may vary in the type and amount of information it holds, usually depending on the type of system. The system makes assumptions about the user generally via some sort of initial on-line questionnaire, which is used to assign a user with a stereotype SM. Performance measures are then collected by the SM and are used to individualise adaptation and output for that particular user. The SM may also store log files, which record input by the user. These may also be used as a performance indicator by the system.

Each AH system also holds a domain model, which describes how the information within the application is structured by describing the relationships between the concepts and nodes etc.

Finally, educational AH systems hold a pedagogical model, which represents knowledge of how to teach based on teaching strategies and diagnostic knowledge.



igure 4.1 Dynamic Model for Adaptive Hypermedia Systems

4.3 User Modelling Issues

To allow adaptation within an AH system even at a very simple level, a model of the user and their progress must be constructed so the system can make adaptation decisions. The first step is to identify the various groups users fall into, for example, experts, occasional users and novices. Having said this it is important to realize that groups do not lie in neat categories, a novice cannot turn into an expert overnight but must work through several continuous phases before they can be classed as an expert. In general, UMs are based on individual user characteristics and should incorporate the fact that the user's capabilities change while working through a session. Therefore, the aim of a SM is also to understand the user's acquisition and retention of knowledge.

4.3.1 SM Attributes – User Characteristics

To initially build a SM the system requires basic information from the user, this is usually obtained via an on-line questionnaire asking about static user characteristics. This information is used by the system to allocate a stereotype SM to the user. Some authors stress the importance of age, cultural background, personality variables and computer literacy levels (Plessis et al., 1985) (Beck et al., 1997) (Spall et al., 1990). Fowler et al. (1986) and Benyon et al. (1988)) suggest a deep user profile including gender, which Beck et al. believe enables a system to customise its interactions to fit the user (although they do insist this is a sensitive issue). Diaper (1987) and Morris (1986) go further to develop a more complex UM including details such as names, spouse, children, home life, jobs, hobbies and leisure interests. Benyon suggests the use of occupation and previous education as well as present mood and personality variables which is felt can influence how users perceive a system and contribute to what they find agreeable/ disagreeable about the system. These attributes may be classed as part of the user's cognitive model holding information such as preferred learning styles, preferred feedback and preferred exercise types (Fowler et al., 1986), Spall et al. (1990), (Elsom-Cook, 1998) (Benyon et al. 1998).

4.3.2 Student Model – Performance Indicators

Student models which are designed for educational AH systems also require data on user performance, which can provide the system with information on how much knowledge the user has gained and their performance during the session. To allow the system to adapt dynamically and provide personalised guidance, performance variables are required, such as those suggested by Cooper (1988), Benyon et al. (1988), Morris et al. (1986) Fowler et al., (1986), Berry et al. (1987) and those proposed for the ADAPTS system (Brusilovsky et al., 1999). Examples include:

- correct and incorrect application usage
- the number and type of requests for help
- time lapses between input
- the number of errors
- the number of documents visited
- how far and the number of times backtracked
- time spent reading documents
- routes taken and the number of times documents were re-visited
- level of competence with system
- history log of previous interaction and domain expertise.

Performance measures, which can also be inferred from logged data, may be used to detect gaps in users knowledge (Zissos et al., 1985). Linton et al. (1999) also suggests that the volume of logged data can indicate the level of the user's expertise. The reasoning behind this is that experts have the ability to plan better solutions as they posses richer cognitive skills resulting in better performance compared with novices who use trial and error and are more likely to take longer routes around an application (Morris, 1986) (Fowler et al., 1986). This has also been Morris' (1986) view, who states that experts are more efficient in their selection strategies and therefore rarely backtrack.

Spall et al. (1990), Seta et al. (1997) and Murphy et al. (1997) present a more detailed UM. Spall et al. (1990) identified two broad types of user models,: conceptual, which is concerned with representing human cognitive processes and quantitative, which deals with user performance data. Spall et al.'s work tested the hypotheses that the user's interactions with an application can provide the basis for a predictive quantitative UM. Spall's experiment did not accurately predict users' expertise based upon correct usage alone even when error times and time lapses were incorporated into the model. Spall et al. stress the difficulty in predicting user expertise from usage alone and recommend that other factors be taken into consideration for example work environment, personal motivation, work pressures and stress. Seta et al. discuss the importance of a static UM with an emphasis on end-user needs. They also state the need for dynamic adaptation which can change individual properties.

4.3.3 Acquisition and Maintenance of User Model Attributes

As stated previously, for the system to make inferences about a user's acquisition of knowledge or expertise it requires certain information, which may be acquired in many ways. Once the information is collected it may be updated during the interactive session to reflect the user's progress or acquisition of new knowledge. Generally, systems use all (Gutkauf et al., 1997) or a combination of the following techniques (Spall et al., 1990) (Fink et al., 1998) (Nill, 1997) and (Morris, 1986).

- Explicit user modelling - a form-style questionnaire, presented to users when they first log on. However, this assumes the user can correctly describe their skill levels (Spall et al., 1990) (Morris et al., 1986). The system then allocates a stereotype SM to each user, which is updated during the session (Benaki et al., 1997).
- Implicit user modelling - the system can make assumptions by analysing, monitoring and observing users' performance with the system and then dynamically update the SM.
- Goal directed modelling - is used to measure the abilities of a user which can not be derived by explicit or implicit user modelling, for example by giving tests (Gutkauf et al., 1997).

In general, student models are hidden from the user. However, providing access to them is seen as vital for gaining the user's trust and acceptance of the system (Strachan et al., 1997). Some applications do display current performance levels as a simple indicator to the student of how well they are doing for example 'See Yourself Write' (Bull, 1997). Some models allow the user to edit their personal SM (Linden et al., 1997). Murphy et al.. (1997) also suggest obtaining information explicitly from the user by asking them directly i.e. the system uses the user to support decision making and information filtering.

4.3.4 Adaptive Assistance and Help

Dynamic assistance is a major issue for adaptive teaching applications. Most AH applications currently adapt the level or type of help given to individual users, for example in Magic box (Plessis et al., 1995) the student has access to dynamic online assistance called Dr. Clue. The help provided is tailored to the specific needs of the individual, by analysing the performance up to that point. For this type of assistance the system requires a richer understanding and information of the users knowledge and beliefs (Ramscar et al., 1997). Hook (1997) suggests three classifications of help;

- User explicitly asks for help
- User asks for guidance
- The system inspects interaction and then warns the user.

Error handling is another aspect of AH systems which needs to be considered. The system needs to decide when an error has occurred, how it should be dealt with and what information should be presented to the user as students prefer positive as well as negative comments (Cohen et al., 1990). The SM can be used to handle errors as Rouet (1992) found that factors such as age, literacy and previous experience were particularly relevant in error handling issues. Errors can be dealt with in the following ways (Morris, 1986):

- by the user with a specific CHANGE command;
- by the system which gives no corrective advice;
- by the system which gives corrective advice when requested;
- by the system which always gives corrective advice;
- by the system which takes automatic corrective advice.

If error handling is computer controlled then decisions need to be made as to whether to interrupt the user or not, or whether to allow the user to set an 'interrupt', where only major errors will be made known to them.

4.3.5 Summary of User Modelling in AH

The concept of a SM is reflected in the traditional classroom environment, where teachers also tend to build mental models of their students (Elsom-Cook, 1988). However, Beck et al. (1997) do not feel that teachers think about individual students at a very deep level of detail, they only use what they need and what is explicitly told to them. In conclusion it is felt that a simpler SM would be more effective than a complex model as there is potentially an infinite number of such models (Elsom-Cook, 1988) (Benyon et al. 1988). However, the use of simple student models rather than complex has the potential for misjudging the variables due to generality of attributes.

SMs use implicit and explicit student modelling to aquire user information and use this model to predict the user's current acquisition and retention of knowledge. The SM must hold static variables such as user characteristics as well as dynamic performance variables.

4.4 A Survey to Construct a Student Model Framework

Due to physical and technical limitations SMs can only capture a limited degree of user characteristics and complex cognitive and even performance processes. Therefore, a SM must be limited to holding only information, which is relevant and representative of the user. This section discusses an in-depth survey, which was carried out using questionnaires, interviews and a focus group. The survey was carried out to establish a generic framework for a SM, which could be used to build stereotype SMs for AH systems. The aim of the study was to address the following issues;

- What variables should be modelled in a SM framework?
- What level of detail is necessary for a SM?
- How closely tailored to individual learners does the model need to be?
- Should the SM be visible to the user?

4.4.1 Survey Techniques

The survey used the following user centred evaluation techniques; questionnaires, interviews and a focus group. Session logging is another useful approach to user modelling (Petrelli et al., 1999) but due to time and resource constraints could not be used. The evaluation was carried out at the University of Southampton, on students who had used Microcosm based learning applications to support their lectures. Users of Microcosm were used in this survey so that some baseline could be established for their responses i.e. all users were answering with respect to one system. Detailed questionnaires were handed out to gain more specific information from the subjects on their computing experience, personality traits and their views on Microcosm (Davis et al., 1992). The questionnaires were handed out by the tutors during their laboratory sessions, in which they used Microcosm based applications designed specifically for their subject area. The questionnaire can be found in Appendix A. Tutors who ran the sessions were given individual interviews. These were carried out to gather information on how tutors felt about the applications they used in their lab sessions and the implications for different types of users. At a later stage, a small focus group was set up comprising eight first year medical students. This opportunity was used to extract more in-depth views on the usability issues of Microcosm and ideas on improvements from the real end users. To gain a wide response many departments were involved in the evaluation.

This study was designed to investigate variables, which could be used in a SM and attempt to filter out those which did not seem to bear any relevance within a university level learning application. Therefore, the general impact of the variables in the questionnaire was measured by using correlation statistics. For example if a significantly high positive correlation existed between the two variables; shy and a novice computer user, it would be assumed that shy users are generally novices at using a computer and therefore this personality trait would be an important variable for the SM. These generalisations were the incorporated into the SM.

4.4.2 Interview Results

Interviews were held to gain an insight into what tutors felt was lacking in the existing CAL applications they used to support their lectures. In general, there was positive feedback on Microcosm as a hypermedia system, most lecturers enjoyed the freedom given to the students by the system. However, some tutors felt that Microcosm allowed too much freedom and the lack of structure did not aid the learning process as it was too easy to get lost. Another problem highlighted was that Microcosm is 'too open' users can open too many windows and become confused. It was suggested that the interface needed improving especially for novices, possibly by restricting the number of windows that can be opened. Tutors also liked Microcosm because of it's ability to maintain various types of data and the functionality presented to the user. One tutor felt that Microcosm was the perfect student aid (given a few changes) and has been using it to teach for over 6 years. On the whole it was felt that the linking features of Microcosm are very useful but that the overall effectiveness of Microcosm depends on the subject matter. It was also suggested that Microcosm lacked specific intelligent aid, which they felt is the way forward for CAL type applications.

4.4.3 Focus Group Results

A focus group was set up which included eight, first year medical students from the University of Southampton. By encouraging an informal discussion, the participants were encouraged to talk freely and openly about their views. On the whole the students enjoyed using the Pathology Application (Kemp, 2000) (authored in Microcosm) and felt it was useful to them as a learning tool especially with a deeper information base. However, they did make comments to help improve the application. One of the main concerns highlighted was the lack of windows management control provided by the system. A recurring problem was that the same image was repeatedly opened because the original window containing the image or text got lost behind overlapping windows. They suggested that the images could be tiled automatically so that more than one could be compared. They also suggested that the application could close windows automatically once the user started to work on the next question. It was suggested that students with a lack of windows management experience could be educated on windows management beforehand. The same problem was found with the multiple choice question box (mcq). Some students preferred the mcq box to remain on top by default but most wanted the choice. It was also suggested that the mcq box could be incorporated into the main text window. Therefore it would not be necessary to display two main windows on the screen all of the time.

Another problem the students found was multiple clicking of the mouse buttons. This mainly occurred because Microcosm does not make it obvious to the user that it is processing the users request as the arrow does not change. Therefore they suggested an 'hour glass' to be displayed each time the system is processing. System feedback to the user is a necessary component of human computer interaction, it is important to keep users informed and confirm that operations are being carried out (Perez-Quinones et al., 1996). It was also evident from what was said that different students wanted different levels of help and assistance. The most important issues to arise from this was that different users had varying expectations from the systems.

4.4.4 Questionnaire Results

This section presents the results from the questionnaire data. The aim was to establish the variables or components required for a student model framework to develop stereotypes. It is important to remind the reader that these results are taken from user's subjective opinions and personal assessment of their experience and skill levels. (The data can be fond in Appendix O).

4.4.4.1 Computing Experience

When user's were asked about their computing experience, it was found that male users are more confident at using computers with 78% of confident computer users being male and only 22% female. This was also the case with e-mail experience, with 73% of experts being male and only 27% female, and the World Wide Web (WWW/ Internet) with 81.6% % of experts being male and only 18.4% female. Similar results were found with word processing and programming experience.

It was also found that users who class themselves as experts at windows management are also experts at using a word processor, MS Office, e-mail and WWW and are generally confident computer users. These users also tended to work with many windows open (many being more than 3) at one time and did not require much detailed help or assistance.

Novice WWW users felt that Microcosm allowed too many windows to be on display at one time which led to confusion, with some experts also agreeing. It was interesting to find that a very high percentage of subjects felt that Microcosm allowed too much windows freedom.

The subjects who felt that they were experts on Windows management generally felt that they could find any relevant documents within the applications when working with Microcosm. These students also did not want the application to offer hints, guidance or change automatically. A higher number of females than males felt that Microcosm windows management allowed too much freedom. This was also the case with users who classed themselves as novice at the following; Microcosm, programming, e-mail, WWW and windows management. These results imply that gender could be a relevant attribute for a SM. General computing experience is also an attribute which could be used by the system to adapt interface type features of an AH system.

4.4.4.2 Computing Experience With Respect to the Various Departments

Analysing the results by department, it was found that the highest percentage of users who classed themselves as experts on e-mail were from the Computer Science Department at 55.6% and the highest percentage of novices were from the Medical Department at 44%. Again the highest percentage of experts on the WWW were from Computer Science at 38.9% with the highest percentage of novices being from the Language Centre 80% and Medicine at 64%. The highest percentage of experts on a word processor were again from the Computer Science Department at 39%, Medical Department at 38% and Archaeology Department at 37% and the highest percentage of novices from the Language Centre and Oceanography Departments. Students from the Computer Science and Mechanical Engineering Departments had the highest number of confident computer users with students from the Archaeology Department being the least confident. Computer Science again had the highest percentage of experts on windows management at 38% with the highest percentage of novices belonging to the Language Centre at 60% and Mechanical Engineering at 58%, Computer Science had the lowest percentage of novices at 5.6%. These results highlighted the differences in computing expertise within various departments of the university, with Computer Science students having the most expertise in most applications and confidence at using computers. These results also show that Medical students seemed to have the highest percentage of novice users for most of the applications. It was interesting to find that although the Mechanical Engineering Department had a high percentage of confident computer users they did not have a high percentage of experts on using the various applications and had a low percentage of windows management experts. These results suggest that this type of information i.e. the main subject/ department background of a user, if held within a SM, could be used to predict a user's confidence and experience levels. These attributes may be used by the system to predict users experience and assist the user while using the AH system.

4.4.4.3 Dynamic Assistance

It was found that both genders needed the same amount of help, although females did find Microcosm significantly harder to use than males. In general, users who are confidant at using computers wanted less help from the application. Users who watch more than 15 hours of television per week, wanted the application to offer assistance and they wanted the ability to ask the application for assistance

The subjects who felt that they could find the relevant documents within the hypermedia application preferred to have the ability to ask the application for assistance but did not require detailed help dynamically.

In general, most of the respondents wanted the option to ask the application for help when they felt they required it. Most of the students preferred assistance in the form of guidance or hints rather than detailed help. Confident computer users did not want the application to provide automatic assistance whereas unconfident users did.

Only 50% of students from the Computer Science Department and 50% from the Medical Department and all the subjects from the other departments wanted the application to make suggestions. These results are significant in that if incorporated into the SM, the system can establish what level of help or assistance a particular user may require.

4.4.4.4 Relevance of Age

Younger users, i.e. those less than twenty one, felt that they were easily distracted while working on a computer but these users also felt that they could find the most relevant documents easily. This outcome implies that there is some significance in the way different age groups work on computers. However, age as an attribute in the SM would require further research, which may prove difficult within a university environment, as on the whole most undergraduate students are under 25.

4.4.4.5 User's Country of High School Education

It was also found that students who had studied at a high school overseas were slightly less confident at using computers (34%) than students who had been to high school in England (45%). This could be a result of language or cultural barriers.

4.4.4.6 Personality Characteristics

At a very general level, it was found that personality characteristics were significant with respect to how users used or perceived a system. For example shy, quiet and pessimistic users tended to find the system had a lack of windows interface control. Again, these types of attributes would require further investigation. However, these results do show that there is some significant difference in attitudes towards how users with various personality traits work using computers.

4.4.4.7 User Information Icons

User were asked which performance indicators they would like to view or have access to on the screen, to help them manage their session while using a CAL application.

<u>Relevant documents not viewed</u>

These documents would be documents within the hypertext application that the system feels the user needs to view for their current goal (i.e. reason for using the application) or skill level. Students from each department and from all the computing confidence levels (novices as high as 88%) felt that this would be useful information, as did those who watch more than 15 hours of television per week, females, novices at e-mail and those who felt that they could not find relevant documents. However, users who work with many windows open at one time (i.e. more than 3) did not require this type of information displayed.

<u>Current skill/ knowledge level with respect to the domain</u>

The students from the Computer Science Department were the only group who did not feel that this would be useful whereas the rest of the subjects did, as did users who spent over 15 hours per week watching television and those who felt that they could find relevant documents.

<u>Number of windows currently open</u>

Users from all computing confidence levels felt that this would be useful information.

<u>Type of errors made</u>

Only Computer Science students did not want information on errors they had made displayed. The rest of the subject groups and unconfident computer users felt that it was important and useful information to be made available to the users of an application.

Time variables such as session length

Indications of time could include the length of the session or time spent viewing a document etc. Again, students from the Computer Science and Medical Departments did not feel that this would be useful information, whereas the remaining subjects did. Generally females, who spent more than 15 hours per week watching television, novices at Microcosm and users who felt that they could find relevant documents also felt this would be useful information.

Documents visited

Regular and confident computer users did not want this information but novices felt that it would be important, as did expert WWW users. Students from all the departments also wanted this information except for Computer Science and Medical students.

These results give some indication of the type of information different types of users would like to have access to while working on a CAL based application.

4.4.5 Survey Discussion

As expected, a higher percentage of males assessed themselves as experts in most of the questions related to computing confidence and application experience. However, a higher percentage of males answered the questionnaires than females. Again as expected, experts on various computer application packages were also experts on windows management and vice versa. Users expert at one application were also more likely to be experts on the others as well. This maybe due to the transference of skills within packages. This indicates that computing experience can be used to give an indication of users experience levels and the type of assistance they would require.

There was a significant difference in computer literacy levels between students from the Science and Arts faculties. It came as no surprise to find that Computer Science had the highest number of expert computer users. The reverse was also true as the Arts and Medicine had the highest percentage of novice computer users. Although users who classed themselves as experts on a word processor did manage to span the subject divide where students from all the faculties felt they were experts. These results indicate that a user's educational background could be a useful attribute for the SM.

User's interests, which many authors include in their user models did not seem to reveal any significant findings. However, this area requires further investigation. Obvious results were observed for students who watch many hours of television or spend a large amount of time using a computer in that the more time a user spends on a computer the more confident they are and the more time a user spends watching television the less confident they are.

The user's background, with respect to country of high school education, did not effect the user's expectations or use of the hypermedia system. Overseas students did not find Microcosm any easier or harder to use or want any more or less help than students who had been to high school in England.

Again, personality variables were interesting but require further investigation. In general personality variables were found not to be highly relevant to the interaction process. However, traits such as shy, pessimistic and quiet were somewhat significant. The user's work/ home environment, number of children, the number of people they live with etc. again did not seem to bear any influence on how the user interacted with or accepted the application. Therefore it was felt that user's background information should not be included in a SM.

It was interesting to find that most users, irrespective of their skill level, wanted detailed information when errors were made and wanted the ability to ask the application for assistance rather than the system making changes automatically.

When asked how much intelligent support users required from an application, all of the three suggestions presented were popular but each with different groups of users. The results of the questionnaire showed, as expected, that computer experts require less help from the system in the form of hints but novices prefer detailed help. Most users felt it was important to be able to ask the application for assistance as required.

There was also a distinct division between the performance indicators that the expert and novice students wanted to be made available. Computing experts only wanted the current skill level and time variables displayed. Whereas novices wanted more information such as documents visited, documents not visited, windows open and various time variables.

4.4.6 The SaDy Student Model

It was evident from the results that many of the variables recommended for use in a SM from the literature review, were not useful in this case and too personal to effect user interaction. However, the results from the survey did reveal some variables, which may be of great importance for building a basic framework for a SM for use with educational adaptive hypermedia applications.

From the results of the survey we proposed the SaDy SM i.e. the Static and Dynamic Student Model shown in Figure 4-2. It is a very basic SM framework for building stereotypes for use in adaptive educational applications, by extracting the variables which could be used by the system to stereotype the user, and variables which could be used to assess user performance and expertise of a user. As the literature revealed there are two distinct areas of user modelling i.e. static and dynamic, where the static holds user characteristic information and the dynamic holds performance type data, which will be dynamically updated while the application is being used. Due to the type of information held in the SM, the variables have been categorised into default, previous and current data. The default data is static and will not change from session to session, the previous data which is also part of the static model contains variables which will be updated during or at the end of a session (this communication is represented by two-way arrows). The current data, which is part of the dynamic model, is used to adapt, assist and guide the user individually. This modular structure of SaDy allows the extraction and insertion of categorised data, which should result in a reusable and portable SM.

4.4.7 Problems With the Study

One of the main problems with this study was that a larger number of students from the Science faculties took part in the survey than from the Arts. It was difficult to investigate the significance of age in great detail as most of the student's ages ranged between 18-24. Again, establishing the relevance of user's interests was difficult to investigate given the time constraints.

4.5 Using The SaDy Student Model in Future Work

As Figure 4-2 shows, the SaDy SM is very complex, and therefore too complicated and difficult to test in a simple study. Also before attempting to analyse and test the model it was necessary to collect relevant data. Therefore, it was decided to run two experiments, one that focused on the static (and some dynamic) variables of the SaDy SM and one that focused on the dynamic variables of the SaDy SM. As a result it was decided that some simple adaptive hypermedia techniques would be used to verify and test the SaDy SM. These experiments are described in the following chapters. For each experiment a pre-session questionnaire was used to obtain information for the static section of the SaDy SM and then a post-session analysis was conducted to refine the SaDy SM.

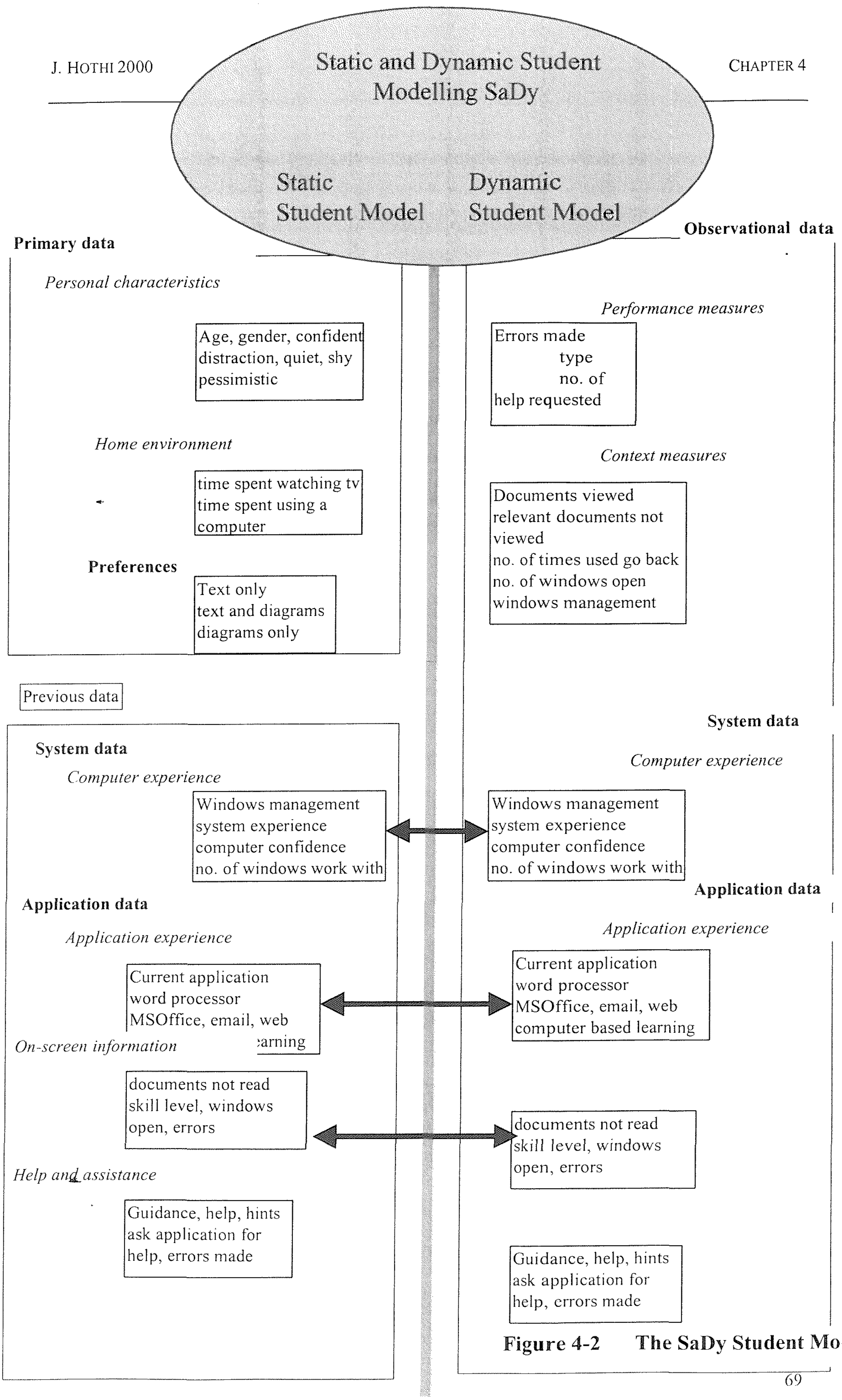


Figure 4-2 The SaDy Student Model

Chapter 5 Experiment 1- Investigating Static Adaptation Techniques

5.1 Introduction

This chapter describes the first empirical study, which was designed to compare the usability of two static adaptation techniques. Two applications were developed, an Adapted Navigation Application and an Adapted Presentation Application. These were compared to an existing (non-adaptive) Archaeology Hypermedia Application. This experiment was also designed to verify part of the static SM derived from the first survey.

5.2 Adapted Hypermedia Applications; Dating Techniques in Archaeology

The material used in this study was provided by the Archaeology Department at the University of Southampton. The documents (and linkbase) was already held in digital format for an existing hypermedia based course in Archaeology. This application was authored in Microcosm by an archaeologist and has been running in the department for the past 3 years. However, only a subset of this application was used for the experiment i.e. Dating Techniques in Archaeology. This existing application was used as the control hypermedia application within this study, i.e. the non-adapted version of the AH applications. The AH applications used the same material and content. The Archaeology Department also provided a subject domain expert who revised the material and the links for the applications. The domain expert also decided what would be considered complex material for novice users.

It was decided that two forms of adaptation would be considered in this experiment resulting in the development of the following applications. (The main differences between each of the applications will be explained in more detail in the following sections).

1. An Adapted Presentation Application
2. An Adapted Navigation Application
3. A Hypermedia Application with no adaptation.

5.3 Implementation

The existing (non-adaptive) hypermedia application was implemented in Microcosm, which was discussed in some detail in Chapter 2. Therefore, it was decided to implement the two AH applications in Microcosm. Microcosm was also chosen because it provides the freedom for storing link information in separate linkbases, which is ideal for building adaptive systems. The applications were presented to users so that they did not require any Microcosm functionality other than highlighted link following and simple windows management, so there was no loss of generality in interpreting the results. As the applications were adapted rather than adaptive, the links were authored at design time.

5.4 Student Modelling Issues

5.4.1 Allocating Generic Student Model Groups

It was decided that the experiment would be kept simple by only supporting two generic SM groups, as in the SATELIT 2 experiment (Akoulchina, 1997); content (or domain) novice and content experts.

The previous study revealed the attributes that might be included within the static section of a SM. It also discussed the acquisition of user information via explicit student modelling. Therefore, in this experiment user information was gathered explicitly by a pre-session questionnaire (which can be found in Appendix D). The results of which were used to allocate each user a generic SM status, i.e. a content expert or a content novice.

In all, there were three sections to the questionnaire; personal characteristics, computing experience (system skill) and content (domain) knowledge. The current domain knowledge was used to allocate each user to one of the SM groups. The personality variables and computing experience were not used to assign SM groups, although this information was used during the analysis stage.

5.5 Application Design

5.5.1 Adapted Presentation Application

The Adapted Presentation Application only supported the shaded fragments of text method of AH. The shaded fragments of text in this application were presented by greying out the background of the text. These shaded fragments of text covered the concept in extra detail e.g. detailed formulae or calculations. This complex shaded information was the same information that was provided to content expert users via a link in the Adapted Navigation Application (see section 5.5.2) (but was hidden from the novice users of the Adapted Navigation Application).

The domain model for the Adapted Presentation Application was the same as the domain model for the existing control Hypermedia Application (see Appendix I). The Adapted Navigation Application had a larger set of documents (for the content expert group) because the complex information was in separate documents. All the users were presented with the same set of documents and links, and therefore only required the use of one linkbase (the standard linkbase).

The user's content knowledge level was used to decide whether a user was told to read the shaded text or to ignore it. Content experts were told to read the shaded text and the content novice users were told not to read the shaded material. Figure 5-1 shows an example of the shading method of adaptive hypermedia that was used in the Adapted Presentation Application. The highlighted links in the application have been underlined for the purpose of this thesis.

5.5.2 Adapted Navigation Application

For the Adapted Navigation Application the user's content skill level was used to decide which of two linkbases would be activated for a particular user i.e. users were assigned to either the content expert linkbase or the content novice linkbase. The content experts were provided with additional links to complex information. (The same complex information that was shaded in the Adapted Presentation Application). In effect the complex information was hidden from (removed from the text for) the content novice users. As a result this application held a slightly larger number of documents (for the content expert group) in the domain model compared to the Adapted

Presentation Application and control Hypermedia Application domain models, due to the complex information that was held in separate documents.

5.5.3 Hypermedia Application – Non-adapted

The literature review (see Chapter 3) revealed that to successfully carry out an evaluation of adaptivity it was necessary to compare the adapted versions of the application to a well designed non-adaptive system. Therefore, the existing non-adapted version of the Archaeology application was used as the control Hypermedia Application. This application had been used successfully by the Archaeology Department for the previous three years for undergraduate students. This control application presented the same links and documents to all types of users. This application held the same set of documents that were used for the Adapted Presentation Application and only used one linkbase (the standard linkbase) i.e. the same domain model as the Adapted Presentation Application.

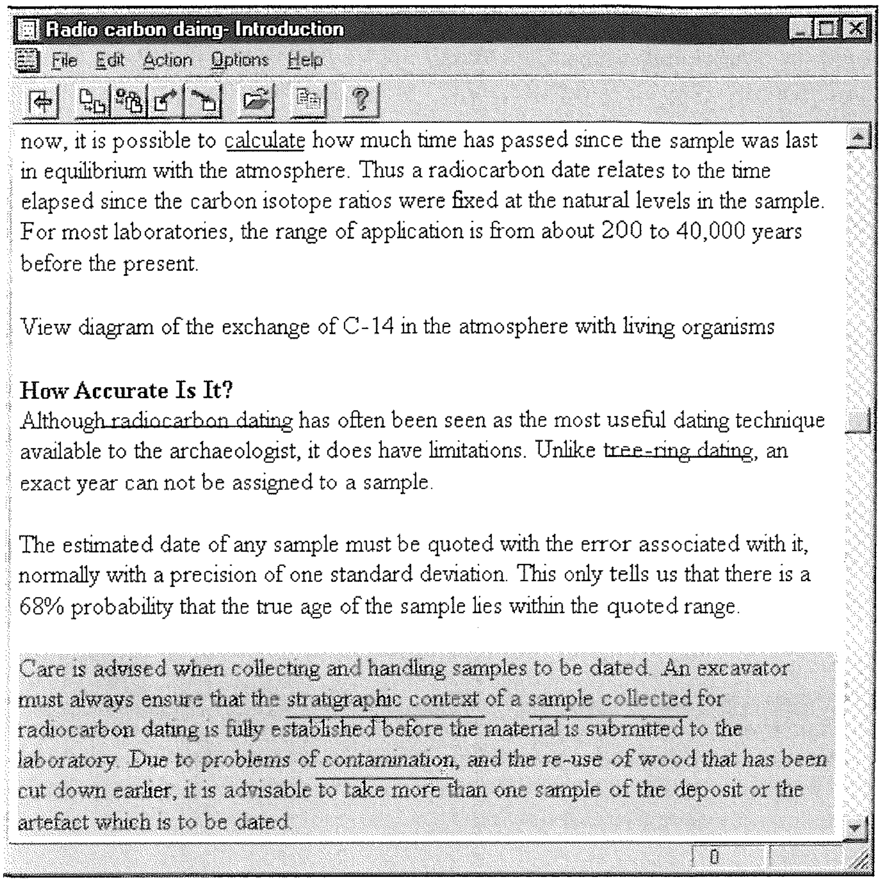


Figure 5-1 Example of the Shading Method of Adaptive Hypermedia

5.5.4 Summary of Design

Table 5-1 below, presents a summary of the differences between the applications, the linkbases that were used for each application and a clearer view of the differences between the two SM user groups.

		Application	
		Adapted Presentation	**Adapted Navigation**
Content skill level	**Expert**	Read shaded text containing complex information	Extra links to complex information
	Novice	Don't read shaded text containing complex information	No links to complex information
Number of linkbases used		1 standard	1 Expert 1 Novice

Table 5-1 Summary of the AH Application Design

5.6 The Pilot Study

A pilot study was set-up where two users carried out a basic evaluation study using the two adapted hypermedia applications. The pilot was set up to investigate the evaluation process and to validate the questionnaires and to establish the amount of time it would take to answer basic questions. Only a sub-set of documents were used in the pilot study, i.e. the radiocarbon dating topic documents. Each user answered a pre-session questionnaire and then followed the instructions to answer the set tasks using each of the three applications in turn. The questionnaire and tasks can be found in Appendix B and C. The pilot study revealed that the tasks took longer than expected and therefore the intended task list for the full scale sessions needed to be shortened. It also revealed that some questions were missing on the questionnaires.

5.7 The Evaluation Study

In total, 19 subjects took part in the experiment, 9 from the Computer Science Department and 10 from the Archaeology Department. Two laboratory sessions were run for the evaluations due to space constraints. The participant's current content knowledge on Dating Techniques in Archaeology and their computing experience varied widely.

The evaluation session started by each user completing the pre-session questionnaire (see Appendix D), from which each user was allocated a SM group, i.e. content novice or content expert. The pre-session questionnaire covered some basic questions on personality characteristics computing skill levels, current knowledge on dating techniques in archaeology and current attitudes towards adaptive hypermedia in educational applications. All the subjects were then instructed to use each application in turn (each subject used each of the three applications during the session). They started off by using the Adapted Navigation Application, followed by the Adapted Presentation Application and finished by using the control Hypermedia Application.

The tasks set were different for each application and the content experts were required to answer further, more specific questions (the tasks for each user group can be found in Appendix E and G). These were designed to encourage the experts to either read the shaded text in the Adapted Presentation Application or to follow the additional links in the Adapted Navigation Application. This also prevented the problem of users increasing their knowledge on tasks completed while using a previous application. However, this still presented a problem for the study in that users felt more comfortable using the applications as they went on to use the second and third. After completing the tasks (which were presented on paper) for one application, the subjects were asked to complete a short questionnaire asking about the usability aspects of that particular application (these can be found in Appendix E and G). This was repeated for the other two applications. At the end of the evaluation session the users were required to complete another more general overall post-session questionnaire reporting on the main differences between each application they used and their overall subjective preferences (the overall post-session questionnaires for each user group can be found in Appendix F and H). The questions on the post-session questionnaires asked the users for their subjective opinions on the adaptation techniques compared to the control application. The results from this experiment are presented in the following chapter.

Chapter 6 Results

6.1 Introduction

This chapter presents the results collected from the first adapted hypermedia evaluation study described in Chapter 5. The results are split into three parts. In section 6.2 we discuss the results with respect to each application to establish the differences if any, between each. In section 6.3 we discuss the data in terms of the two generic SM groups, to establish the differences in performance between each group. In section 6.4 we discuss the subjective preferences of the subjects regarding adaptation techniques in CAL. The discussion and the summary of the findings is presented in section 6.5.

6.2 Comparing Individual User Performance By Application

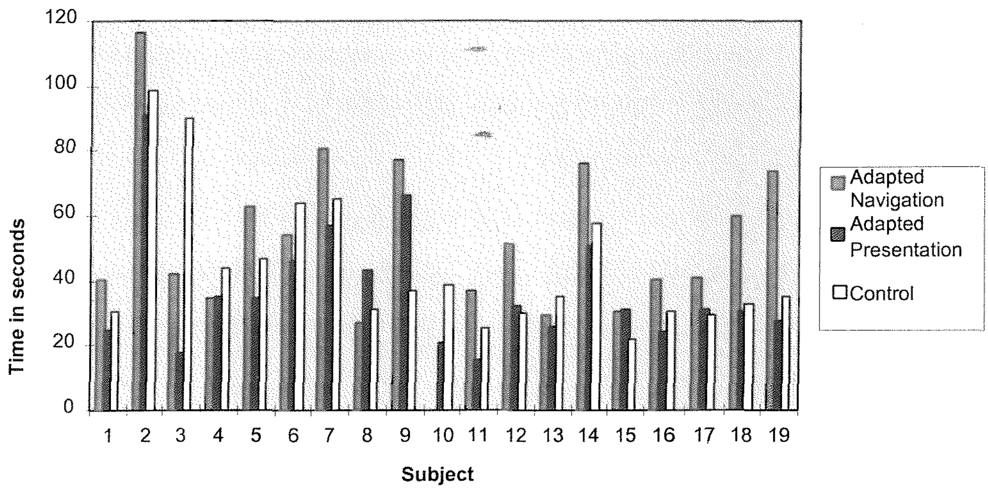


Figure 6-1 Average Time Spent Per Document Within Each Application

6.2.1 Average Time Spent Per Document Within Each Application

Figure 6-1 shows the average time users spent viewing each node for each application. The graph clearly shows that a high percentage of subjects spent a large amount of time viewing each node while using the Adapted Navigation Application. The subjects also spent a significantly greater amount of time visiting each node within the control Hypermedia Application. A high percentage of subjects spent only a brief time on each node within the Adapted Presentation Application. These results are interesting, because although users spent less time completing the tasks for the Adapted Presentation Application, Figure 6-2 clearly shows that the percentage of tasks answered correctly was also higher for the Adaptive Presentation Application compared to the other two applications. The content novice users performed the worst on the control application compared to the results form the AH applications.

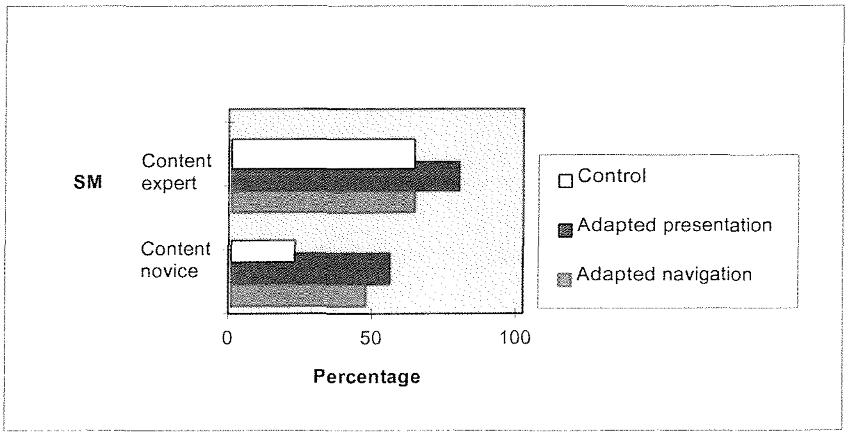


Figure 6-2 Percentage of Correctly Completed Tasks Within Each SM Group for Each Application

6.2.2 Total Session Time For Each Application

The total time that each subject took to complete the tasks for each application were analysed. No obvious pattern was observed, each student completed each session in different times. However, on the whole the subjects tended to complete the Adapted Presentation Application in less time compared with the other two applications.

6.2.3 Average Time Spent Viewing Relevant Documents Within Each Application

An analysis of the average time users spent viewing relevant nodes within each application was also carried out. The relevant nodes that were taken into account included all nodes containing information required to answer the set tasks. On average, most of the subjects spent less time viewing the relevant nodes while using the Adapted Presentation Application compared to the other two applications.

6.3 Analysis of Data With Respect To Each Generic SM Group

In this section, the results are presented by showing comparisons between the two SM groups, i.e. content expert and content novice. The results shown are an average of each SM group.

6.3.1 Average Time Spent Viewing Documents Within Each Application For Each Student Model Group

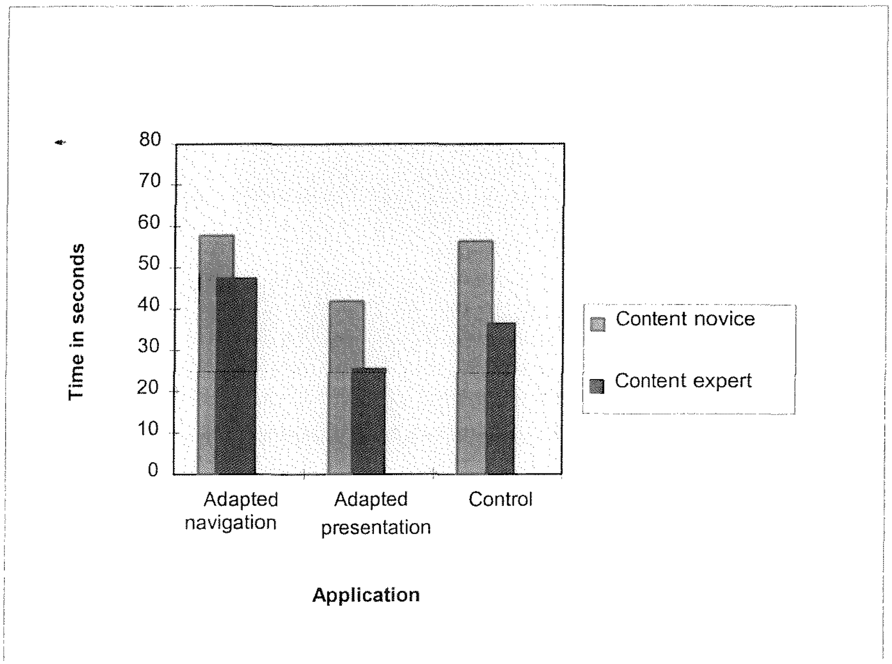


Figure 6-3 Average Time Spent Viewing Documents Within Each Application for Each SM Group

Figure 6-3 shows the average time that each SM group spent per document for each application. The graph clearly shows that for both the SM groups, the least amount of time spent per document was within the Adapted Presentation Application. The graph also highlights that the content expert group spent the least amount of time per document within each application but the least in the Adapted Presentation Application. It is also clear from this graph that the content novice group spent more time per document within each application, implying that they needed to spend more time to understand the content. However, they did spend the least amount of time on each node within the Adapted Presentation Application compared to the other two applications, which show similar results.

6.3.2 Average Time Spent Per Relevant Document Within Each Application For Each Student Model Group

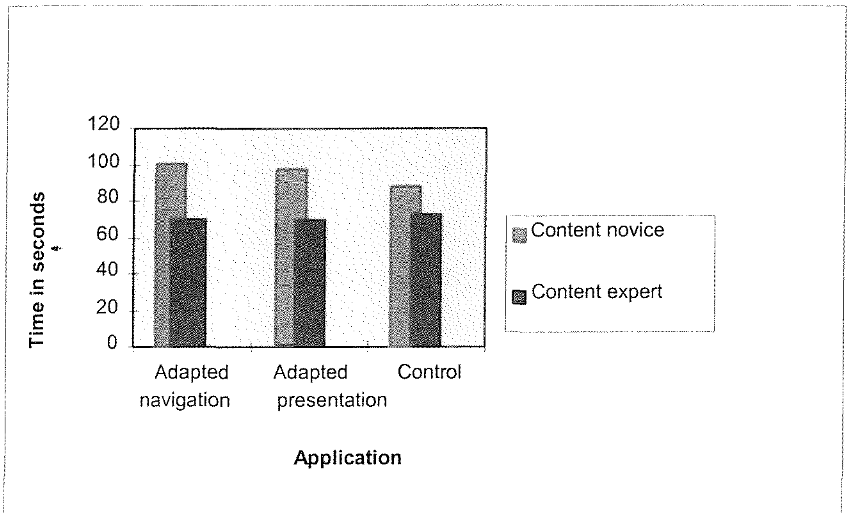


Figure 6-4 Average Time Spent Viewing Relevant Documents Within Each Application for Each SM Group

Figure 6-4 shows the average time that each SM group spent on relevant documents (i.e. nodes which contained the relevant material to answer the tasks) within each application. This graph shows that each group spent a similar amount of time viewing relevant documents while using each application. Overall, the content expert group spent less time viewing relevant nodes compared to the content novice group. The content novice group also spent the highest amount of time viewing relevant document within the control Hypermedia Application and the Adapted Navigation Applications.

6.3.3 Average Session Completion Time For Each Application For Each Student Model Group

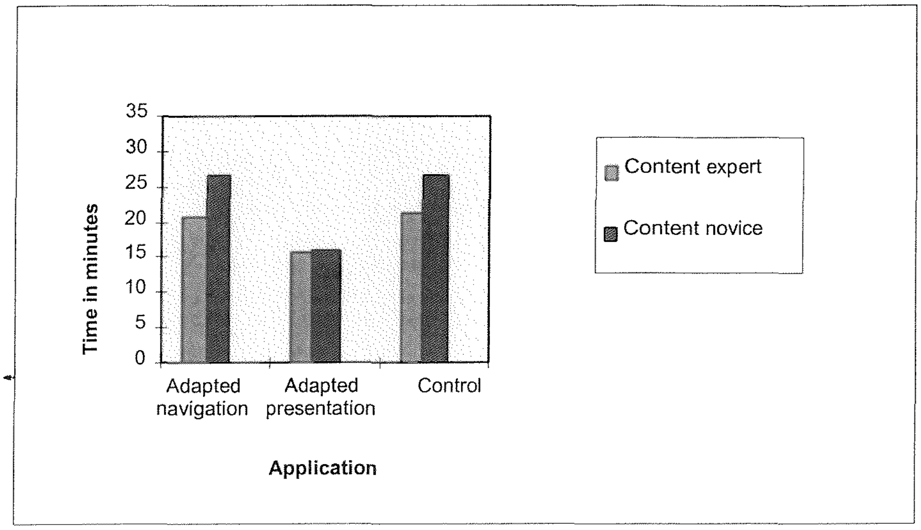


Figure 6-5 Average Session Time Within Each Application for Each SM Group

Figure 6-5 shows the average time each SM group spent completing the tasks within each application. The graph clearly shows that both groups completed the Adapted Presentation Application in the least amount of time compared with the other two applications. On the whole, content experts completed the applications quicker than the content novice group, except for the Adapted Presentation Application where similar completion times were observed.

6.3.4 Average Number of Nodes Visited Within Each Application For Each Student Model Group

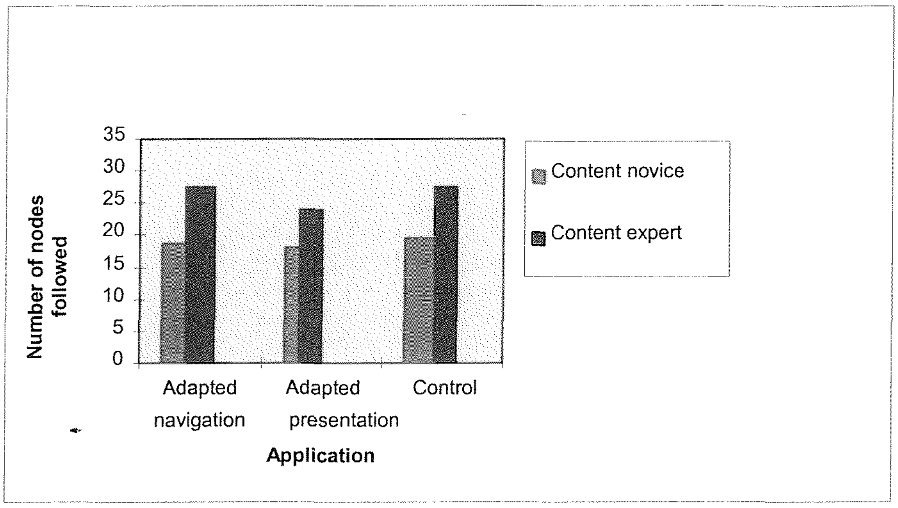


Figure 6-6 The Average Number Of Nodes Visited Within Each Application for Each SM Group

Figure 6-6 shows the average number of nodes that each SM group visited within each application. Again, the graph shows that each SM group visited a similar number of nodes within each application. Both SM groups when using the Adapted Presentation Application visited the least number of nodes. The content expert user group tended to visit more nodes than the content novice SM group within each application.

6.4 Subjective Preferences on Adaptation Techniques

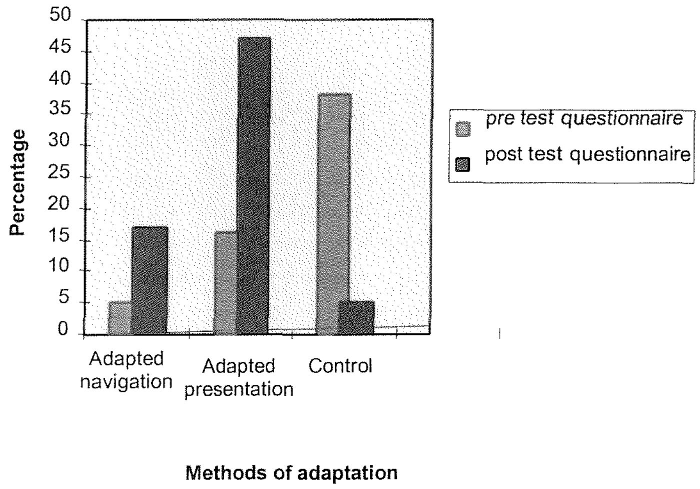


Figure 6-7 Subjective Preferences for Adaptation Techniques

Figure 6-7 clearly shows the differences between the pre and post-session questionnaires. The graph highlights the shifts in preferences towards the acceptance of adaptation within educational hypermedia applications. In the pre-session questionnaire, a high percentage of subjects felt they would be more comfortable and prefer to use conventional hypermedia applications and a very low percentage preferred the use of AH. However, the post-session questionnaire presents a significant shift in preference towards the Adapted Presentation Application with the acceptance of the Adapted Navigation Application also increasing. Some of the subjects also suggested that it would be useful to use several techniques in combination.

6.5 Discussion

This section of the report presents an analysis of the results gathered from the first evaluation study. The discussion is divided into three parts. The first section talks about the use and relevance of the static SM within the scope of the experiment. The second section goes on to investigate the differences between the various forms of adaptation techniques that were used in the study. Finally section 6.6 concludes and summarises the findings from the evaluation.

6.5.1 Verification of the Static Student Model

Even though the initial questionnaire was basic, it did appear to place each subject into the correct SM group. This can be determined from the results where subjects within each SM group exhibited similar working patterns. Ideally, it is necessary to obtain in-depth knowledge of the user's content and system skill levels, possibly via an on-line questionnaire.

A correlation analysis was carried out on the data, which proved to be a very useful measurement technique. For this analysis, each variable (e.g. computing confidence, e-mail expertise, personality variables etc) was correlated against each of the other variables in turn, to show which variables were significant to the way the subjects worked and how they used the applications.

It is important to remind the reader of the results from the SaDy study that found personality variables and gender relevant to how a user interacts with a system. There is also a general acceptance within the user modelling community (Diaper, 1987), (Beck et al., 1997), (Spall et al., 1990) that these are relevant variables for constructing a basic framework for a student SM. However, within the scope of this study, personality variables were not used to assign SM groups and were only used at the analysis stage i.e. to test if users with certain personality traits performed in similar ways. The analysis of these results showed only some positive correlation. The lack of a high percentage of positive results from the current analysis could be attributed to the small population sample that took part in the study as opposed to the larger sample, which took part in the first SaDy study.

In general the results did show that pessimistic users tended to work with a fewer number of open windows on screen at any one time. It was also found that shy users preferred an educational application to offer as much context sensitive help as possible and present a combination of adaptation techniques within one application than less shy users.

It was interesting to find that the number of hours a subject watched television and (as expected) the number of hours a subject used a computer during a week was very significant in how a user accepted and worked with a system. The results showed that subjects who spent more than 15 hours a week watching television were generally novice at using a word processor and users who spent less than 15 hours per week watching television were experts. Similar results were found when analysing windows management experience. There was also a high positive correlation between users who spent less than 15 hours per week watching television and their expertise on the WWW and e-mail. These subjects also worked with more than three windows on the screen at any one time. Again, users who spent less than 15 hours per week watching television felt that they had easily found the relevant documents within each application. It was also found that these subjects visited a fewer number of nodes per application and completed each set of tasks faster. As expected, the subjects who spent more than 15 hours a week using a computer were experts on most packages, worked with more than three windows on the screen at any one time and felt that they were more experienced in windows management.

Again, as expected, the number of windows the subjects worked with at any one time also had some relevance to their experience on other packages. This also effected the way they worked during the evaluation session. That is the more windows a user works with at one time the more experienced they were on most of the packages listed in the pre-session questionnaire. Also the subjects, who worked with more than three windows open at any one time, in general managed to complete the set tasks in a shorter period of time. Windows management experience e.g. resizing, moving etc. also proved to be highly relevant to how well a particular user interacted with the system. Experts were also experts on most of the packages i.e. word processors, e-mail and the WWW. Also, as expected, there was also a high positive correlation between the number of windows a user works with and their windows management experience.

It was also interesting (but again expected) to find that there was a strong relationship between windows management experience and the subject the user was studying. A high proportion of the Archaeology students were novices at windows management whereas Computer Science students were experts. These results mirror the results obtained from the first survey.

6.5.2 Dynamic Student Model Attributes

As the graphs in this chapter show, the analysis produced some interesting results and gave a good insight into which variables can be used to determine user's content and computer expertise dynamically. It is accepted that the use of time to assess or predict user performance is a grey area as it is difficult to assess what a user is doing within a particular timeframe without technology such as eye trackers etc. However, many authors also feel that the time between mouse clicks can be relevant (Boyle et al., 1991) and others agree that in general novices take longer finding target items (Young et al., 1990). This study also found that the time spent viewing a document could be used as an indicator of the level of content knowledge, since the content expert SM group spent less time per document compared with the content novice groups. However, the average time spent on relevant documents by each SM group proved to be a more significant variable for predicting knowledge level than the average time spent on each document. The content novice SM group spent more time viewing relevant documents than the content expert SM group. These results were as expected due to the difference between the knowledge levels of each SM group. That is, the content novice group was expected to spend more time on each relevant document as they needed to learn and absorb even the most basic material, whereas the content expert SM group either already knew the basics or were recalling what they had learnt previously. The content expert SM group also completed the study in less time compared to content novices.

The number of nodes visited is another variable that was found to be a significant performance indicator e.g. to establish the user's content knowledge levels, where the content expert SM group visited a higher number of nodes within each application compared with the content novice SM group.

In summary, during the analysis it became obvious that members of each SM group worked in a similar manner. The number of links they followed, the time they spent on each document and the total session times etc. were similar to other members of the same group.

An obvious explanation for subjects spending, on average, a greater amount of time viewing each node within the Adapted Navigation Application could be due to the fact that this application was the first one the subjects were asked to use. So, this application was used to absorb the material and create mental maps of the application structure. The results from Graph 6.1 are significant, in that although on average the subjects spent less time per node within the Adapted Presentation Application the percentage of correctly answered questions was higher for this application in comparison to the other two applications.

6.5.3 Adaptive Techniques in Hypermedia

Although all the subjects had used some form of hypermedia, AH was a new concept to most of the subjects who took part in the evaluation study. Therefore, when they were asked in the pre-session questionnaire; *which method of adapted hypermedia do you prefer?* A high percentage of the subjects chose the conventional hypermedia application. The reason stated for this choice was that they preferred to have access to all the material and all the links available within an application. However, after the evaluation session they were asked the same question in the post-session questionnaire and their views had dramatically shifted. A very low percentage liked the idea of the control Hypermedia Application and a very high percentage preferred the use of the Adapted Presentation Application with the Adapted Navigation Application also becoming more appealing once it had been used. Some of the subjects even proposed the implementation of a combination of adaptation methods within one application. Again the reasons stated for these choices were that users could still have access to all the material but could avoid wasting time reading material, which may be too complex or irrelevant for their particular task. Also all the information would be available in case they wanted to know more about a particular topic, i.e. no information was hidden from them. Other statements given included;

- help focus on relevant material instead of wasting time shifting through large amounts of text
- avoid getting lost, and
- it was easier to find relevant material.

These views were reflected in the answers given to other questions within the questionnaire; 78% of the subjects felt that all the material within an application should be made available to all types of users, 'just like a conventional book', and 63% agreed that shading out fragments of irrelevant text reduced information overload.

In summary, the subjects spent the least amount of time (on average) on documents and on relevant documents within the Adapted Presentation Application. The difference between the session completion times for each application was not significantly different but on the whole it was the lowest for the Adapted Presentation Application. An important finding was that each group correctly answered the highest number of questions when using the Adapted Presentation Application.

6.6 Conclusion

Before concluding the results from this study it should be noted that due to the small number of participants in the empirical study, the SaDy SM has not been revised since the original diagram in Figure 4-2. However, in conclusion we can begin to construct a set of variables for the SaDy SM that could be significant in predicting and analysing user performance. The static SM could contain information such as the number of windows a user works with, their general windows management, e-mail and WWW experience and the time spent watching television and time spent using a computer per week. These variables are important because it shows how computer literate a user is (generally) and therefore provides some indication of how users interact with a system as well as indicating the user's navigational capabilities. The user's educational background, e.g. the subject the user has studied/ is studying, can also reflect user experience. Personality traits such as shy, pessimistic etc. and gender (in combination with other variables) can also affect the way a user perceives a system.

The dynamic section of SaDy SM could include variables such as, the number of nodes visited, the average time spent viewing nodes, task completion time and the average time spent on relevant nodes etc.

It is important to note here, that the users cognitive model of the application(s) is a separate and more complicated issue and this area was deliberately left out from this work. The user's knowledge and mental maps of the application domain are very important in building a full and detailed model of how a user interacts with a system. However, the aim of this work was only to concentrate on looking into the variables, required to build basic SM, which may be used for educational AH and how that information could be used to predict user behaviour.

This chapter presented a preliminary experiment using static AH methods in a study, which was targeted to investigate the usability of the shaded fragments of text technique of AH. To develop the shaded AH technique further it was necessary to carry out another study, which combined adaptive presentation techniques with an adaptive link technique.

The next chapter will discuss the design and implementation of a second empirical experiment, which compared the usability of two AH techniques, shaded text and dynamic ANS. Again, as explained at the end of Chapter 4, the SaDy SM was not implemented in the study due to its complexity. However, data was gathered via performance logs etc. to carry out a post-session analysis of the dynamic SaDy SM variables.

Chapter 7 Experiment 2- Comparing Dynamic Adaptive Techniques

7.1 Aim of the Empirical Study

This chapter discusses the aim of the second study which was twofold; firstly to investigate the usability and effectiveness of combining two AH techniques within one AH application, and secondly to further explore the attributes of the SaDy Student Model.

7.2 Comparing AH Methods

The survey presented in Chapter 4 introduced the variables of the SaDy SM. The first experiment, which was described in Chapter 5, introduced the shading method of adaptive presentation where the results showed that users positively accepted this method of AH in educational applications. The aim of the second experiment, which is discussed in this chapter was to further develop and investigate the usability and effectiveness of the shading method of AH by combining it with a dynamic adaptive navigational technique of AH. When considering the design of the AH applications it was important to consider the following aims of AH applications in general:

1. To reduce information overload
2. To reduce cognitive overload
3. To provide dynamic navigational support
4. To allow all users (with any knowledge level) access to all the available material

The literature review in Chapter 3 covered the AH techniques studied within the field of AH and highlighted the advantages and disadvantages of each technique. This experiment was designed to draw on these studies to improve existing, or combinations of complementary techniques.

The previous experiment (discussed in Chapter 5) found that the AH method of shading fragments of text was a novel yet practical addition to the AH taxonomy. Shading can be viewed as a form of stretch-text without highlighted links i.e. a form of adaptable hypermedia. It also results in reduced cognitive overload with a reduction in the number of visible links. However, this technique alone only proposed a solution to issues 1, 4 and, to some extent 2 from the list above. So it was necessary to introduce dynamic adaptive adaptation method(s) to approach issues 2 and 3. It was decided that these could best be solved using ANS techniques such as link annotation such as those used in studies by Brusilovsky et al. (1995) and De Bra et al. (1998a). The link annotation method results in a slight increase in the user's cognitive load, but it is still the most attractive method of ANS as users are presented with all available links and nothing is hidden from them. Link hiding was not used as some studies found that users did not like the hidden aspect and spent their time running the mouse over words to search for hidden links (Da Silva et al., 1998).

This experiment used the material from the previous empirical study i.e. Dating Techniques in Archaeology. Three applications were decided upon for the following reasons. It was necessary to develop one application that included the shaded method of AH and this application also needed to be tested in combination with an ANS method of AH to produce a complete AH application. As discussed above the ANS method chosen was one that had been proven to be acceptable by users in other studies i.e. link annotation (Brusilovsky et al., 1998a). To validate the experiment it was necessary to compare the application that combined the shaded method of AH with link annotation with one that supported the same form of link annotation alone. Again, a control application was required to make valid comparisons against, therefore the existing Archaeology Hypermedia Application was used as the control.

7.3 Designing the Applications

The empirical study was designed to compare the usability, effectiveness and acceptance of AH. Each application is discussed below.

7.3.1 Shading and ANS AH Application (A)

Application A was an AH application that supported two methods of adaptation; that of shading (a static adapted form of adaptation) and dynamic link annotation (a dynamic adaptive form of adaptation). Although, the shaded method of adaptation in this experiment is referred to as adaptive, the shading of text was implemented statically at design time it was therefore not a truly adaptive form of hypermedia. However, this method can easily be made adaptive by using a similar approach to adaptive - conditional text systems. I.e. where fragments of text can be shaded dynamically and presented to users as they progress though an application depending on the users current skill level.

Within this experiment, each document had between zero and 3 sections of shaded text. All users were presented with the same documents and links. Content expert users were told to read the shaded text as they would be ready to understand the content and the novice users were told to ignore it as it would be too detailed for their current knowledge levels, although, they could read it if they so wished. ANS was provided in the form of link annotation by changing the colour of links presented to individual users. The method of ANS used within this study was adapted from Brusilovsky's study (1994), which used the traffic light metaphor for colouring the links depending on one of four states (this is discussed further in section 7.4.1).

7.3.2 ANS AH Application (B)

Application B was the second AH application and only supported the ANS form of AH. The method of link annotation was implemented in the same way as for Application A. Within this application, text that contained complex information or formulae etc. was placed in a separate document and the user was required to follow a link to view this material. (This was the same text that was shaded in Application A). Content expert users were told to follow more complex information type links as they would be ready to understand the content and the novice users were told to ignore them as it would be too detailed for their current knowledge levels. However, they could follow these links if they felt it would help them. Again, users were presented with the same documents and links.

7.3.3 Hypermedia Application – Non-adapted (C)

The Hypermedia Application, Application C was used as the control application with no adaptation. Although this application was based on the original Hypermedia Application used in the previous study, (as in application B) complex information and formulae were presented in a separate document and the users were required to follow a link to view this material. All users were presented with the same documents and links. Each type of link was presented in blue regardless of the state i.e. visited, not-visited etc.

7.4 Structure of the Domain Model

7.4.1 ANS - Link Types

The ANS was only supported in Applications A and B. All the links were coloured blue within Application C and therefore were not distinguishable.

Link types	Status of the Link	Colour of the Link
Not visited – (relevant) link	Ready-to-be-learned	Dark blue
Visited – link	Learnt	Dark blue and underlined
Not relevant link	Not-ready-to-be-learned	Magenta
Not visited - generic link	(always ready) Ready	Light blue
Visited – generic link	Learnt generic links	Light blue and underlined
'More complex information' - Not visited – relevant link	Ready-to be learned	Dark Blue
'More complex information'- Not relevant link	Not-ready-to-be-learned	Magenta
'More complex information'- Visited link	Learnt	Dark blue and underlined

Table 7-1 ANS Code for Applications A and B

Table 7-1 shows the different types of links and their respective colours. The ready-to-be-learned links were coloured dark blue. Each document within Applications A and B contained a number of these links. As a user visited a document, the system assumed they were ready to view other documents within the same concept cluster (CC are explained in more detail in section 7.4.2.1). As explained previously, Applications B and C included a special case of 'more complex information' type links, which were links to complex information on the current document. (In Application A this more complex information is given as shaded text in the same document). This extra information was reached by following a link, which read 'for further information follow this link'. For users of Application B, these 'more complex information' links were classed and presented in the same way as the concept type links on the current document. For example, if the user had followed a not-ready-to-be-learned link, i.e. a magenta coloured link, these 'more complex information' type links on that document were also coloured magenta. If the user followed a ready-to-be-learned link, i.e. a dark blue link, the 'more complex information' type links on the document were also coloured in dark blue and once visited these remained dark blue but were also underlined in the same colour.

The set of not-ready-to-be-learned highlighted links were coloured magenta to show the user that they were not ready for these documents. Reference or generic links, were light blue links to topic summaries, glossary terms and images. Again, once visited, these links remained light blue but were underlined in the same colour (i.e. all visited links were underlined). Finally, at the end of each node was a list of contents for that particular topic, this is discussed in section 7.4.2.3.

All possible links within the hypertext network were functional and users could follow any link, even the not-ready-to-be-learned magenta coloured links. However, if this type of link was followed the system would pop up a dialogue box on the screen reminding the user that they are not ready for this particular document and should therefore return to the previous document. Once the user clicked 'OK' on this dialogue box the document would still be displayed and it would be up to the user if they closed the document and returned to the previous one or continued to read the new document. At this point the ANS was disabled until the user returned to the correct source document (i.e. the one from which they followed the not-ready-to-be-learned link). Although all the links were functional, users were advised to traverse only the dark blue (concept) links and the light blue (generic) links. The idea behind the ANS was only to support and aid users, it was entirely down to individual users if they did or did not follow the advice provided by the system.

In summary the link annotation mechanism used the following algorithm:

1) Follow link
2) Display destination document
3) System accesses the database to establish the status of the links to be displayed
4) If the link followed is a ready-to-be-learned link display all links with relevant colours and underline link to source document to show it has been visited
5) If the link followed was a not-ready-to-be-learned link pop up dialogue box stating that the user is not ready to view this information and should return to

previous document and freeze the status of all links until the user has returned to the source document.

In effect, the application held an implicit model of the user where the system knew which documents the user had visited within each concept cluster and which were not visited.

7.4.2 Domain Model

Any AH system must have complete knowledge of the set of relevant documents for each concept and have knowledge of which concepts need to be learned before others i.e. a complete representation of the domain model. The structure of the domain model was based on the existing Archaeology Hypermedia Application. The concepts were divided into Concept Clusters (explained below) and each concept was presented within a set of (one or more) documents. The overall domain model can be found in Appendix I.

7.4.2.1 Concept Clusters (CC)

The domain model was developed by placing all the Dating Techniques in Archaeology documents within one of the five sub-headings in Dating Techniques in Archaeology, i.e. Radiocarbon Dating, Hydration Layer Dating, Thermoluminescaence, Dendrochronology and Microscopy. Each topic was then split into four Concept Clusters (CC), which resulted in 4 levels of content detail i.e. CC 1-Basic Concepts, CC2-More Information Concepts, CC3-Detailed Concepts and CC4-Complex/ Difficult Information Concepts. This is shown in Figure 7-1. Finally the individual concepts that comprised each Concept Cluster were presented within one or more documents. These documents were a combination of the text documents and the three types of generic documents. These are described in section 7.4.2.2. The structure described in Figure 7-1 was used for Application A. Applications B and C had a similar domain model as Application A but also had additional links to 'more complex information' type documents (whereas for Application A, this same information was left in the document but shaded out).

Basically the user needed to first learn the concepts in CC1 before the system will allow them to visit documents in CC2 and again the user must complete and learn CC2 before the system would allow them to visit CC3. This was also the case for CC4.

7.4.2.2 Generic Type Documents

Three types of generic reference documents were designed: Type A which gave a brief summary of the current topic, Type B which were glossary type documents and Type C which were images and diagrams. All the generic documents were reached only via generic type links. The generic links to these documents were distinguished firstly (for Applications A and B) by the colour of the link i.e. light blue, and secondly the difference between the types of generic links was made explicit in the document text itself (for all the applications). That is, links to Type A generic links were distinguished by the topic name being highlighted. For example, 'Radiocarbon Dating' was highlighted in light blue, which indicated that this link would lead to a summary of this particular topic. Type B generic links were distinguished because glossary type terms such as 'carbon' were highlighted in light blue. Finally, links to images etc were distinguished because the text read that 'to see a diagram of ......follow this link'. At the start of the session the users were informed about the differences between each type of generic link. However, for users of Application C all links were the same colour i.e. blue, including generic links.

7.4.2.3 Tables of Contents as Hypertext Overviews.

Many of the studies in the reviewed literature talked about the importance of preventing users becoming disorientated by providing some navigational assistance. Some systems use graphical views as a navigational aid but this approach would only work for small hypertext networks. For a highly connective network, graphical displays would lead to information overload (Conklin, 1987) and it would be impossible to display on-screen. Some studies on the use of hypertext have shown that browsing through a table of contents is a preferred approach over other types of methods (Capmpagnioni et al., 1989). Therefore it was decided to include a linked contents list at the bottom of each document. This content list only referred to documents within that particular topic area.

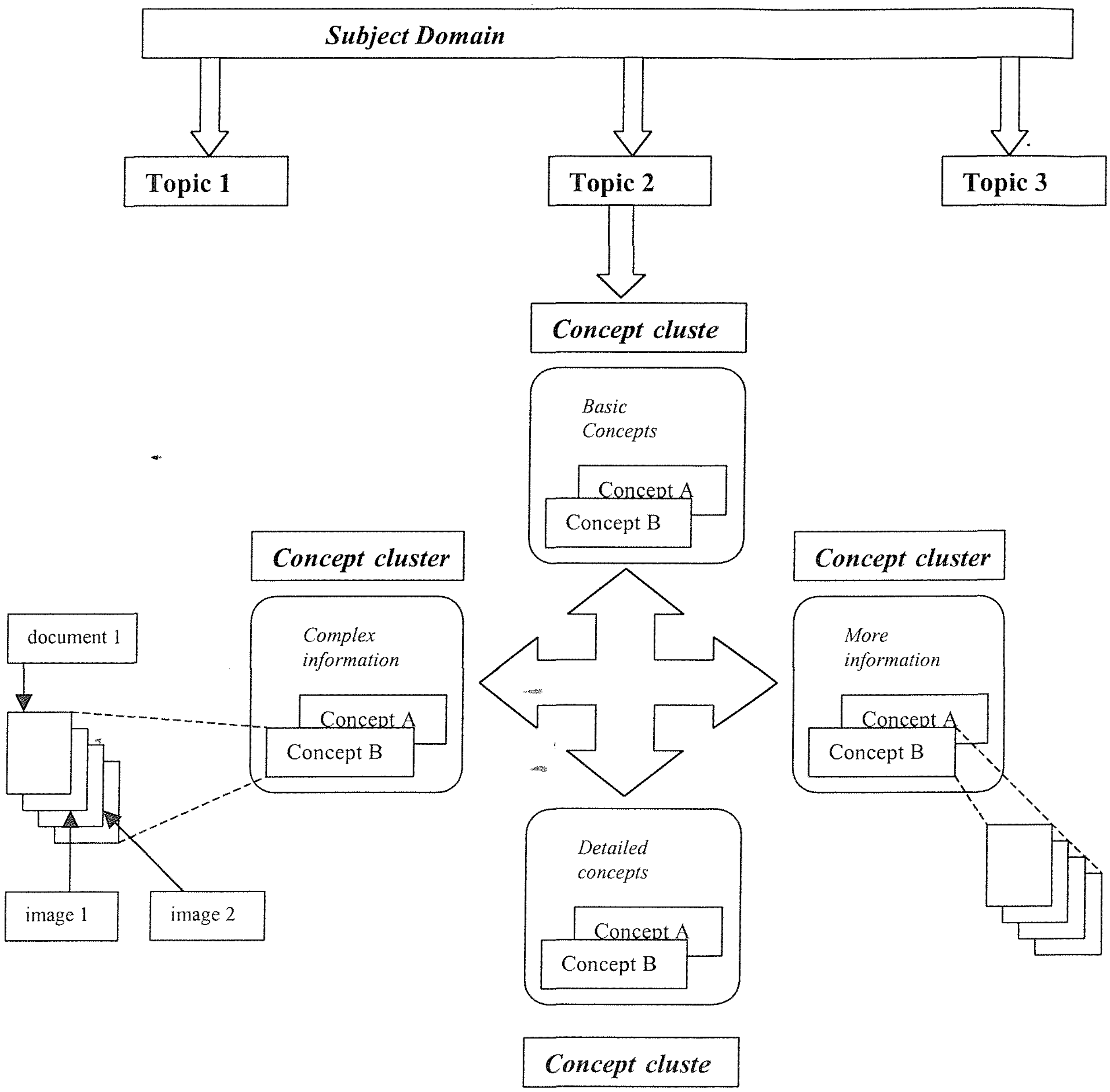


Figure 7-1 The Domain Model

7.5 Implementing the Applications

Applications A and B were designed to provide ANS in the form of link annotation. This was achieved by dynamically changing the colour of links and underlining all visited links. Only one linkbase was used for each application implying that all users were presented with the same links. This was simply because dynamic ANS was used to individualise the status of links to each user.

7.5.1 Developing a Filter for Microcosm

Again, Microcosm (Davis et al., 1992) was used to develop the applications. This hypermedia system was chosen for several reasons but the most relevant being the simplicity of authoring hyperlinks. For a detailed discussion on Microcosm see Chapter 2. The only functionality of Microcosm that was used in the applications was the use of highlighted links and simple windows management i.e. as in the WWW. Therefore, there was no loss of generality when using the Microcosm Hypermedia System in the study. Another important point to note here is that in Microcosm generic links are not highlighted, however, for the purpose of this study, they were.

To support the AH applications it was necessary to design and implement a new filter for Microcosm. The aim of this filter was to support the link annotation for the ANS in the AH applications. The filter annotated the links in two ways;

1. dynamically change the colour of links
2. underline visited links

Dr. Ian Heath from the Multimedia Lab developed all of the coding for this filter at the University of Southampton. The algorithm for changing the colour of links was described in section 7.4.1.

A Microsoft Access database was created, which held all the relevant information for each document set within both the AH Applications A and B. The database held the title of each document, the unique identifier, etc. It also represented the relationships between each document and which Concept Cluster they were in. The domain structure was explained in section 7.4.2. This database was used by Microcosm to make decisions on the current status (colour) of each link depending on which document the user had already visited.

7.5.2 The Student Model

The advantages of modelling the user's knowledge have been the subject of much debate over the years. The literature review revealed two opposing views; the model builders who feel that detailed models are essential in individualising applications which although powerful are quite complex (Orwant, 1991). On the other hand we have the model breakers who question the benefits of using complex SMs and the benefit of them to the teaching process (Sandberg, 1987). The literature review on the evaluation of AH applications revealed that most studies to date only support simple SMs which represent stereotypes and not individuals (Brusilovsky, 1998), (Da Silva, 1998), (De Bra et al., 1998a), (Kaplan et al., 1993). Most studies tend towards a simple approach on the grounds that even if it is partially successful, it will facilitate the evolution and adoption of more powerful modelling techniques.

As discussed at the end of Chapter 4, it was decided that the proposed SaDy SM was too complicated to test in a small study, so only a basic SM was used in this experiment. However, investigating the SaDy Student Model was still an objective of the overall project and was investigated in a post-study performance based analysis. Relevant information was collected during the evaluation sessions by the questionnaires and log files and analysed at the end of the study.

7.5.2.1 Establishing User Knowledge

Some systems assume that learning has taken place by the fact that the user has visited a document (or by passing a test). In most studies, knowledge in AH systems is considered 'learnt' by the user reading/ visiting a document (Da Silva, 1998) (Trella et al., 2000). Therefore, although no SM was held explicitly for the users by the system, the AH applications assumed the user had 'learnt' once they visited a document and 'not learnt' if they had not visited a document. Also the system assumed that the user had not learnt if a document was visited via a not-ready-to-be-learned link.

Although not used within this study, another method that can be used by the SM to assess user's progress is the notion of time. Time in a SM is a very important concept and was investigated during the analysis of the experiment.

7.6 The Evaluation Study

7.6.1 Running The Study

In all, 32 participants took part in the study, 22 male and 10 female. The subjects were a mix of post-graduate students from the Archaeology and Computer Science Departments at the University of Southampton. The pre-session questionnaire and questions revealed (as expected) that all the Archaeology students were classed as content (domain) experts and all the Computer Science students as content novices. Content (domain) level within this experiment refers to the user's knowledge level on the subject domain i.e. Dating Techniques in Archaeology. It is also worth pointing out that all the users had some experience of using computers, i.e. they all had some windows management skills but were not all necessarily system experts.

The participants who took part in the study were randomly allocated to one of the three applications. This was done by listing each participant and randomly assigning 11 names to Application A, 11 to Application B and the remaining participants were assigned to Application C. A week before the first evaluation session was due to start the participants from the Archaeology Department were given a short lecture on how to use Microcosm and a brief introduction to the aims of the research. The participants were told about the set up of the study and what was required of them and were also given the pre-session tasks (see Appendix J) to answer on paper (these tasks were the same as they received during the evaluation session). The tasks were designed in a multiple-choice question (MCQ) format to make it easier to analyse the answers. The participants were also given the pre-session questionnaire (see Appendix K) and were asked to complete the tasks and pre-session questionnaire during this preparation session. The Computer Science students were already familiar with Microcosm and therefore did not required a lecture on the use of Microcosm. However, the aims of the research were explained in detail. They were also given the pre-session questionnaire and the same set of tasks to answer as the Archaeology students. Again, they were required to complete and return these before their evaluation session.

Each user arranged a separate time to visit the lab and carry out the session. As the sessions were carried out on an individual basis, it was possible to videotape each session. At the start of the session the participants were given the same set of tasks as the pre-session tasks (again, presented on paper). However, this time the aim was to use the application to help answer the questions during the sessions.

During the session the participants were encouraged to 'think aloud' i.e. talk through what they were thinking as they used the application. The evaluator was present throughout each session to video tape the session and in case of problems. The participants were also encouraged to ask questions if they encountered problems. Each user was presented with the same first page on Dating Techniques in Archaeology. This was simply an introductory page with a linked contents list containing an entry for each main topic. Subjects were able to answer the tasks in any order they wished and were directed to navigate using only the highlighted links. No user, not even the computer scientists attempted to use any of the Microcosm functionality. At the end of the evaluation session each participant was given a post-session questionnaire Some of these questions were the same as the questions in the pre-session questionnaire and were added to test if subjective attitudes had shifted following the session. The rest of the questions concentrated on the attitudes towards the application itself. Although the pre-session questionnaire was the same for each participant the post-session questionnaires differed slightly and were tailored to each of the three applications. The questionnaire designed for Application A users can be found in Appendix L. The questionnaire designed for Application B users can be found in Appendix M, and The questionnaire designed for Application c users can be found in Appendix N.

7.6.2 Evaluation Techniques

The literature review revealed that evaluation studies carried out within the field of AH are scarce. The studies that have been carried out are mostly informal or used a small number of subjects. This study was aimed at presenting a detailed investigation into the usability and the effectiveness of AH methods within the field of education, by using a structured approach to evaluation.

Figure 7-2, which has been adapted from (Nielsen, 1990) shows the various areas of usability which may be studied and the dashed arrows show the areas of usability that were investigated within this study (with respect to the AH applications and not the hypermedia system). The usability attributes for hypertext systems are explained in more detail in Chapter 3. The literature also showed that authors used conventional qualitative and quantitative evaluation methods to study AH applications. The evaluation techniques used in this study were pre and post-session questionnaires, observation techniques using a video camera, software logs, and think aloud techniques. To measure subjective satisfaction, users were asked directly using the pre and post-session questionnaires. When filming the video camera was set up to capture the user's face, reactions and the computer screen. It was also possible to establish a precise time log using the video recordings. The performance logs were collected automatically by the hypermedia system. The logged data showed individual trails through the application for each user.

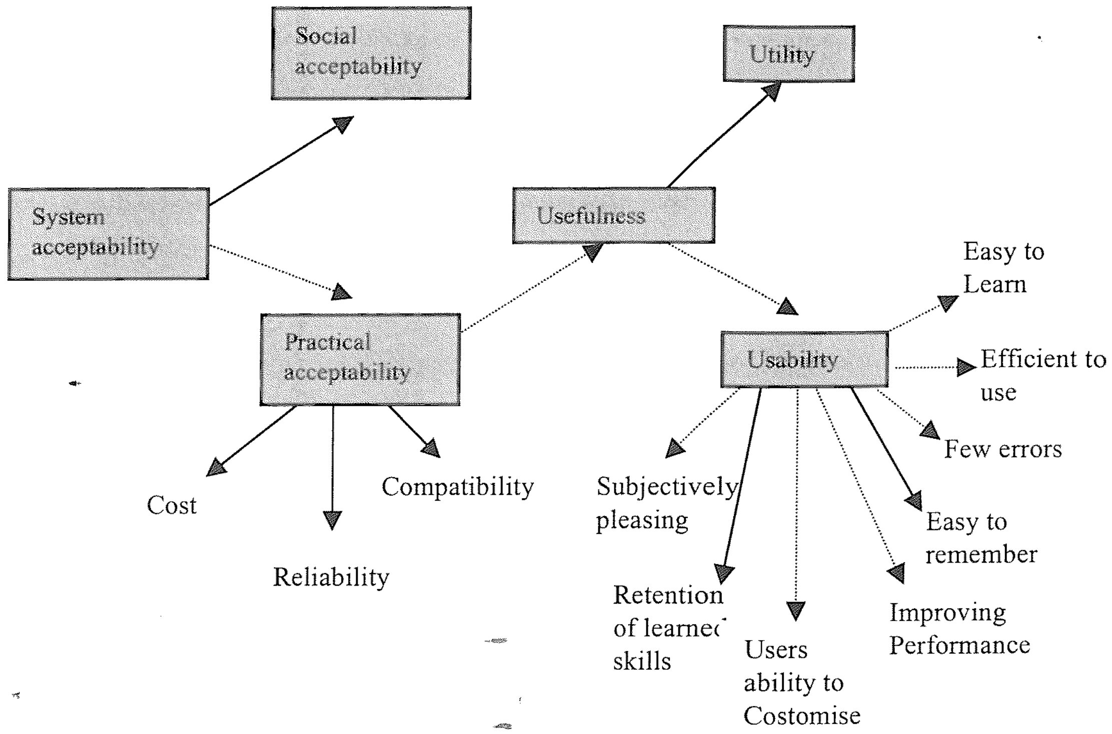


Figure 7-2 Usability Testing

7.7 Conclusion

This experiment involved the comparison of three applications. Two AH applications were developed one supporting ANS and shading and the other only supporting ANS. The third application was used as a control Hypermedia Application without AH techniques. The evaluation study used a combination of techniques to gather information in an effort to investigate the usability, acceptance and effectiveness of the different types of applications. The following chapter presents the results obtained from this study.

Chapter 8 Results From Experiment 2

8.1 Introduction

The results presented in this chapter are divided into 2 main parts. The first section presents the results for the main usability areas, which were investigated during the study; subjective opinions, learnability, ease of use and flexibility. This section is followed by an analysis of user performance data. A discussion and overall conclusion follow these two sections.

The applications will be referred to as Application A, B and C with respect to the following:

- Application A= the AH application with ANS (in the form of dynamically changing the colour of the highlighted links) and the shading of fragments of text
- Application B= the AH application with ANS (in the form of dynamically changing the colour of the highlighted links) only
- Application C= the Control Application which was a hypermedia application with no AH techniques.

8.1.1 Statistical Tests

Table 8-1 below shows the tests which were performed on each set of data

Sets of data analysed and compared	**Measure**	**Type of data**	**Relationship of analysed data**	**Test of significance difference**	**Test of correlation**	**Measure of average**
Comparison of pre-post questionnaire results for each participant	Non-parametric	Ordinal	repeated	Wilcoxon	Spearman	Median
3 sets of application usage data	Parametric	Ranked	independent	ANOVA	Pearson	Mean
3 sets of application questionnaire data	Non-parametric	Ordinal	Independent	Kruskal-Wallis H	Spearman	Median

Table 8-1 Statistical Tests

Pearson's and Spearman's tests of correlation describe the relationship between two variables. Both produce a correlation coefficient as a result. Spearman's correlation is used for parametric data (ordinal scale) and Pearson's correlation is used for non-parametric data (interval scale). The Wilcoxon test is applied to repeated measures or matched pairs. This test was used to examine the differences between the participants' scores in the pre and post-session questionnaires. The result of the test investigates the direction and the size of the difference. The Kruskal Wallis test is similar to the Mann Whitney test but compares the results of two or more groups. The ANOVA test was used to compare the difference in the means of the results. This test was also followed by the independent t-test to establish where the differences lay.

Most of the tests were carried out at the 5% significance level (i.e. probability of 0.05) unless otherwise stated. Some of the tests presented a hypothesis (Ho), where the null hypothesis is stated as Hn.

8.2 Subjective Opinions

8.2.1 Attitudes towards the Use of Computers in Education

As the literature review revealed, it was important to assess users' views on the provision of dynamic assistance or help provided by a learning application. In the pre-session questionnaire users were asked if they preferred to ask the application for assistance when required; 90% of respondents agreed that the user should ask for assistance, 3% were not sure and 6% disagreed. Only 48.5% felt that the application should offer assistance dynamically, 27% were not sure and 24% disagreed. Users were asked, if intelligent tutoring systems would be more beneficial if both options were available to assist them. Only 33% agreed, 15% were not sure and 51.5% disagreed.

8.2.2 Requirement for Having Tutors Present While Using CAL

When asked if tutors should be present while students use CAL applications, 66% from the pre-session questionnaire agreed. However, this decreased to 53% after the evaluation session. The percentage that were not sure increased from 18% to 30% and the percentage of users who disagreed remained roughly the same.

Ho= There is a difference in the pre and post study attitudes of the subjects towards the presence of a tutor while using a CAL package.

Hn= There is no difference in the subjects attitudes toward the presence of a tutor while using a CAL package before and after the study

Using the Wilcoxon test for significance of difference, at the 5% significance level we can say that the result of p=0.02 is statistically significant. Therefore we can reject the null hypothesis and conclude that there is a difference between the subject's pre and post-session attitudes towards a tutor being present while using a CAL application.

About half of each content skill level group still felt that a tutor should still be present while students used CAL based applications.

8.2.3 Comparing CAL and Lectures

When asked if users enjoyed using a computer when studying, 84% agreed and 15% were not sure. However, no user disagreed. In the post-session questionnaire users were asked if they preferred CAL applications to paper based learning methods. There was an equal split where 31% agreed, 33% were not sure and 33% disagreed. These results were broken down into each application user group which can be seen in Table 8-2.

I preferred CAL to paper based learning methods	Agree	Disagree	Not sure
Application A	5	2	4
Application B	2	7	2
Application C	2	2	6

Table 8-2 Preference of Cal to Paper Based Learning Methods

When asked if CAL packages can complement lectures, the figures remained roughly the same pre and post-session. The percentage agreeing increased by 4% and the percentage that were not sure reduced in the post-session response by 6%, from 9% to 3%. All respondents felt that CAL could complement lectures. When users were asked if CAL packages could be used instead of lectures, again the results remained the same pre and post-session, with only 18% of the subjects agreeing, 27% not sure and 51% disagreeing.

When asked if users preferred lectures to CAL, 4 users from Application A agreed, 5 were not sure and 2 disagreed. 8 Application B users said they preferred lectures to CAL and 3 were not sure. For Application C users only 1 said they preferred lectures to CAL, 6 were not sure and 3 disagreed. Out of the users of Applications A and B, who found the ANS confusing, 47% said that they preferred lectures to CAL, 39% were not sure if they preferred lectures to CAL and 13% said that they did not prefer lectures to

CAL. There was also a great difference between the skill level groups where 80% of content experts and only 17% of content novice users preferred lectures to CAL.

8.2.4 Attitudes to the Access of Information while Studying

Users were asked in the pre and post-session questionnaires whether information should be hidden from the user if they are not yet ready for it. In the pre-session questionnaire, only 23% agreed, 27% were not sure and 49% disagreed.

The pre-session questionnaire also asked if users felt that the advantage of using books was partly because they could access any part of the text, 75% agreed, 21% were not sure and only 3% disagreed. Respondents were also asked if they preferred lectures because they were only exposed to information that they were ready for, 21% agreed, 24% were not sure and 54% disagreed.

8.2.5 Acceptance of Adaptive Hypermedia

Participants were asked in the pre and post-session questionnaires whether AH is a beneficial concept in supporting CAL applications. 57.6% agreed in the pre-session questionnaire, this increased to 66.7% in the post-session questionnaire. 33% were not sure if AH could be useful, this percentage decreased to 15.2% in the post-session questionnaire. These results are shown more clearly in Figure 8-1, where it can be seen that all users felt that AH was a beneficial concept.

Ho= There is a difference in the pre and post-session attitudes towards the acceptance of AH

Hn= There is no difference in the pre and post-session attitudes towards the acceptance of AH

Using the Wilcoxon test for significance we can say that the result of $p=0.03$ is statistically significant. Therefore we can reject the null hypothesis and conclude that there is a significant difference in the pre and post-session attitudes towards the acceptance of AH. Also a high percentage of content experts (86%) and content novice (80%) agreed that AH is a useful concept.

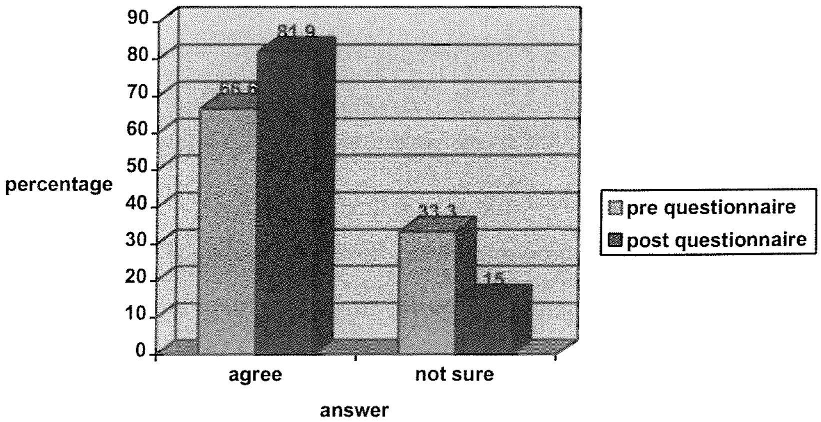


Figure 8-1 AH is a Beneficial Concept

In the post-session questionnaire users were asked if they enjoyed using their particular application, 7 Application A, 7 Application B and only 4 Application C users agreed. When all users were asked if they would consider using their application as a CAL package, 81.8% agreed, 12% were not sure and 3% disagreed. Again by breaking these results down into each application user group, 10 Application A users, 8 Application B users and 9 Application C users felt that they would like to use their respective applications again. No user said that they would not use the application again.

8.3 Learnability

8.3.1 Finding Relevant Information

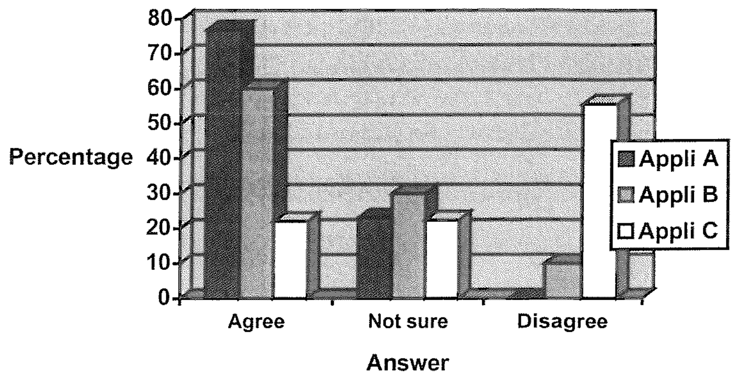


Figure 8-2 Could find The Information Quickly and Easily

All the users were asked if they could find the relevant information to answer the tasks quickly and easily while using the application. The results are shown in Figure 8-2, where it can be seen that the highest percentage that agreed was from the Application A user group. The highest percentage to disagree was from the Application C user group. To test if the users were answering the questions as truthfully as possible this question was asked again but re-phrased and similar results were found. Table 8-3 presents some general statistics. The results ranged from 1 (agree) to 5 (disagree). The mean was the lowest for Application A implying that most users agreed that they found information quickly and easily while using Application A.

Ho= There is a difference between the user groups of each application in their opinions on whether they found relevant information quickly and easily.

Hn= There is no difference between the user groups of each application in their opinions on whether they found relevant information quickly and easily.

The result from the Kruskal Wallis test (p=0.006) was significant and therefore we can reject the null hypothesis and accept that there is a difference between the opinions of the user groups of each application in their opinions on whether they found relevant information quickly and easily.

I found information quickly and easily	Mean	Min value	Max value
Application A	2.1	1	3
Application B	2.5	2	4
Application C	3.3	2	4

Table 8-3 Ease of Finding Information

Of the users of Applications A and B who found the ANS confusing, 50% felt that they could find the information quickly and easily and 50% were not sure. Out of those that did not find the ANS confusing, 68% felt that they could find the relevant information quickly and easily, 25% were not sure and 6% felt that they did not find information quickly and easily.

Out of the Application A user group, 7 out of the 11 users who all found the shading useful also said that they found the relevant information quickly and easily. Even the users who did not find the shading useful still felt that they could find the information quickly and easily.

60% of content experts felt that they could find the information quickly and easily compared to only 52% of content novice users.

8.3.2 Re-visiting Nodes

Users were asked in the post-session questionnaire if they needed to revisit documents to complete the set tasks, 8 Application A, 10 Application B users and 10 Application C users agreed. This is discussed further in the performance analysis in section 8.6.

8.3.3 Acquisition of Knowledge

In the post-session questionnaire users were asked 2 questions relating to their acquisition of knowledge on Dating Techniques in Archaeology (DTA). Firstly they were asked if they had picked up any general knowledge on DTA and secondly they were asked if they had learnt the basics in DTA. Out of the users from Application A, 8 agreed with the first question. However, all agreed that they had learnt the basics. For users of Application B only 5 agreed with the first statement, 5 were not sure and 1 disagreed. However, 9 agreed that they had learnt the basics after using the application and 2 were not sure. Only 6 Application C users felt that they had learnt a little after using the application and 3 were not sure. However, 8 felt that they had learnt the basics of DTA and 2 said that they were not sure.

Ho= There is a difference in each user group's subjective opinion on their acquisition of basic DTA knowledge between each application.

Hn= There is no difference in each user group's subjective opinion on their acquisition of basic DTA knowledge between each application.

Using the Kruskal-Wallis test the result of p=0.35 was found not to be statistically significant. Therefore, we must accept the null hypothesis and conclude that there is no difference between each user group's subjective opinion on their acquisition of basic DTA knowledge between each application.

8.4 Ease of Use

8.4.1 The Application Was Confusing to Use

Users were asked if they found their particular application confusing to use. 2 users from Application A agreed. 8 users from Application A, all 11 from Application B and 7 users from Application C disagreed. This question was repeated but phrased differently to test if users were answering the questions as truthfully as possible. Again similar results were found.

Out of the Application A user group, all those who found the shading of fragments of text useful did not feel that the application was confusing to use.

8.4.2 AH Techniques

It was interesting to establish if users of Applications A and B felt that their application contained too many AH methods in one application. Only one user from the Application A user group agreed with this statement. 3 from Application A and 3 from Application B were not sure. 7 users from Application A and 8 from Application B disagreed. By combining these two groups, half of the users who found the application confusing to use felt that the application maintained too many AH techniques.

Out of the Application A user group, those who felt that shading of fragments of text was useful, did not feel that Application A supported too many AH techniques.

And out of those who felt that the shading was not useful, half were not sure if the application supported too many AH techniques and half disagreed.

Of the Application A and B user groups who found the ANS confusing, half felt that the application supported too many AH techniques.

8.4.3 Feeling Lost and Disoriented

All users were asked if they became lost and disorientated while using the application. Only one user from each application agreed. 9 users from the Application A user group, 9 from Application B and 9 from Application C disagreed.

Ho= There is a difference in each application user group's subjective opinion if they got lost and disoriented while using their particular application

Hn= There not a difference in each application user group's subjective opinion if they got lost and disoriented while using their particular application

Using the Kruskal-Wallis test the result of p=0.87 is not statistically significant therefore we must accept the null hypothesis and conclude that there is not a significant difference in these results.

8.4.4 Usefulness of the Table of Contents

All users were asked if the table of contents was a useful navigation aid. All the users from the Application A user group agreed, 8 Application B users agreed with 1 not sure and 2 disagreeing. 9 Application C users also agreed with only 1 disagreeing. When asked if the table of contents provided some idea of the structure of the hypertext network, nearly 100% of users agreed.

8.4.5 Understanding the Adaptability of the Applications

When asked if understanding the adaptability of the system was easy, 8 users from Application A and 8 from Application B agreed.

Ho= There is difference between each user group in the understanding the applications.

Hn= There was no difference between each user group in the understanding the applications.

The Mann Whitney test (p=0.8) showed that the difference between each user group was not significant and we can therefore accept the null hypothesis.

7 out of 11 Application A users felt that the shading method of AH was easy to understand and also useful. Out of the Application A and B users, 50% of those who found the ANS confusing, felt they could understand the adaptability features of the system and 50% were not sure. Out of those who did not find the ANS confusing to use, 81% felt that they understood the adaptability and 18% were not sure and no one disagreed. Table 8-4 shows that a higher percentage of Application A users felt that ANS was useful compared with the users of Application B.

Application	The guidance given by the coloured links was useful	The guidance given by the shaded text was useful
A	10	7
B	7	-

Table 8-4 Guidance Provided by the Applications

8.5 Flexibility

8.5.1 Shading of Text Fragments

Users of Application A were asked if the shading of the fragments of text disrupted the reading flow, only 1 user agreed and the remaining 10 disagreed. For those users from the Application A user group who found the application confusing to use, half felt that the guidance provided by shading was useful while the other half were not sure. Out of those who did not find the application confusing to use most agreed that the guidance was useful.

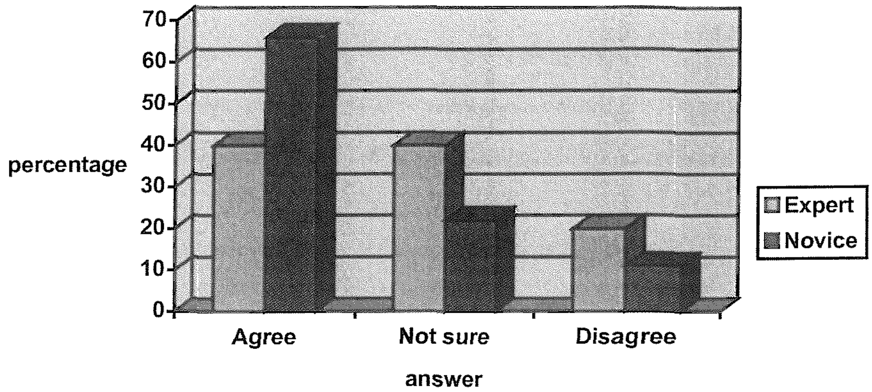


Figure 8-3 The Guidance Provided by the Shading was Useful by Content Knowledge Level

The Application A user group was asked if the shading of fragments of text was useful, 8 agreed, 2 were not sure with 1 disagreeing. Figure 8-3 shows that a higher percentage of novices agreed that the shading was useful compared to content experts.

8.5.2 Users Feeling in Control of The Application

All users were asked, if they felt in control of the application they used. 8 Application A, 9 Application B and only 6 Application C users agreed. This can be seen in more detail in Table 8-5.

Application	I felt that I had control over the system			
	Agree	**Not sure**	**Disagree**	**Total**
A	8	2	1	11
B	9	1	1	11
C	6	4	0	10

Table 8-5 Control Over the System

Out of those who found the application they used confusing, half felt that they did not have control over the system. Out of those who did not find the application confusing to use, 84% did feel they had control over the system, 11% were not sure and 3% did not feel in control. Out of the Application A user group, those who felt that the shaded text was useful, all felt they had control over the system. Those from Application B who found the application confusing to use did not feel that they had control over the system. Out of the users who did not find Application B confusing to use, 9 felt that they had control over the system.

8.5.3 Guidance Provided By the Applications

All users were asked if the guidance provided by their particular application was useful. The results for Applications A and B are shown in Figure 8-4. 9 Application A users agreed compared to only 5 from Application B and 2 from Application C. Half of the Application C users disagreed.

Of the Application A and B user groups who did not find the ANS confusing most felt that the guidance provided was useful. Out of all those who did not find the ANS confusing, nearly three-quarters said they enjoyed using the application. Of those who felt that the ANS was confusing, half felt that the guidance provided by the application was useful and half-disagreed.

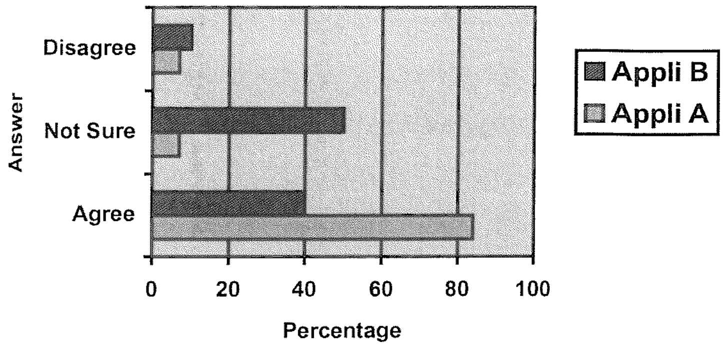


Figure 8-4 Guidance Provided by the Application Was Useful

8.5.4 The Guidance Provided by the ANS

8.5.4.1 Guidance of ANS

The users of Applications A and B were asked if the ANS led to confusion. Only 1 Application A user and no one using Application B agreed. 8 Application A users and 8 Application B users disagreed. Of the Application A user group those who felt that the shaded text was useful also felt that the ANS did not lead to confusion. In fact, all the users of Application A agreed that the guidance provided by the ANS was useful.

Out of the users of Application A and B, only 11% of content experts and 14% of novices felt that the ANS led to confusion, i.e. 70% of experts and 64% of novice did not find it confusing. The results were found to be similar for both sets of users.

Users of Applications A and B were also asked if the guidance provided by the ANS was useful. The results are presented in Figure 8-5. 10 users of Application A and 7 from Application B agreed. Out of the Applications A and B user groups, even those who found their application confusing to use agreed that the guidance provided by the ANS was useful. Table 8-6 shows some relevant statistics. Out of the Application A and B user groups, 89% of content expert users and 71% of content novice users felt that the guidance provided by ANS was useful.

The guidance provided by the ANS was useful	SD	Median	mean	Min value	Max value
Application A	2	0.49	1.9	1	3
Application B	2	0.84	2.4	1	4

Table 8-6 Guidance Provided by the ANS

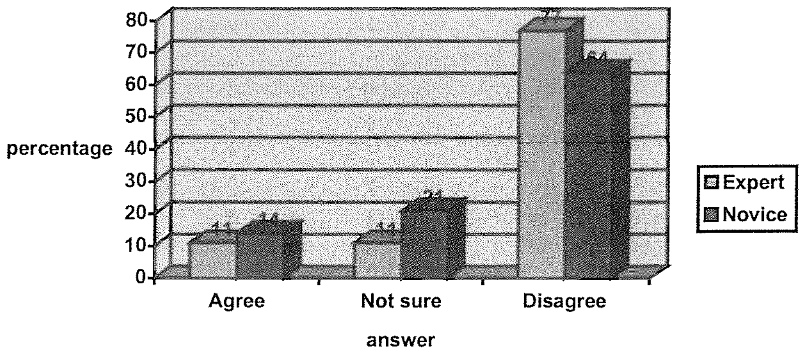


Figure 8-5 Guidance Provided by ANS Led to Confusion by Content Knowledge Level

8.5.4.2 Distinguishing Between Various Types of Links

8 users from Application A and 8 from Application C said that they found the links within the text useful with 1 not sure. Again, out of the users who felt that their application was confusing to use, all felt that the links within the text were useful.

Of the Application A and B user groups, half of those that felt the ANS led to confusion agreed that they could distinguish links they had already visited. Out of those that did not find the ANS confusing, 93% felt they could distinguish visited links and 7% disagreed. Out of the Application A user group, 9 said that they could distinguish links they had visited. These results were the same when users were asked if they could distinguish between the not-ready-to-be-learned and reference type links. Even out of the users who felt that their application was confusing to use, all said that it was easy to distinguish links they had already visited. These results are presented in detail in Table 8-7 which highlights similar results for Applications A and B compared to Application C. Table 8-8 presents some general statistics, which also show the difference between the AH applications and Application C.

The next set of hypotheses and results determine if any differences existed between each (of all three) application user groups when distinguishing the different types of links.

Ho= there was a difference between users being able to distinguish visited links within each application.
Hn= there was no difference between users being able to distinguish visited links within each application.

The Kruskal Wallis test showed that the result was significant p=0.00. Therefore we can reject the null hypothesis and accept that there was a difference between the user groups in distinguishing visited links.

Ho= there was a difference between users being able to distinguish ready-to-be-learned links within each application.
Hn= there was no difference between users being able to distinguish ready-to-be-learned links within each application.

The Kruskal Wallis test again showed that the result was significant at p=0.00. Therefore we can reject the null hypothesis and accept that there was a difference between the user groups in distinguishing ready-to-be-learned links.

Ho= there was a difference between users being able to distinguish reference type links within each application.
Hn= there was no difference between users being able to distinguish reference type links within each application.

The Kruskal Wallis test showed that the result was significant again at p=0.001. Therefore we can reject the null hypothesis and accept that there was a difference between the user groups in distinguishing reference type links.

Due to the significant results found in the Kruskal Wallis tests it was necessary to compare each user group result with that of the other two groups to establish where the differences lay. Therefore, the Mann Whitney test was carried out comparing Applications A and B, A and C, and B and C. The results showed that no significant difference existed between the results of Applications A and B. A highly significant difference was found between Applications A and C, and B and C. This implied that there was a significant difference between the results from the AH applications compared to Application C.

Remarkably, 100% of the content expert users felt they knew which links to follow next whereas only 52% of content novice users agreed with this statement. This could be due to the fact that the experts had prior knowledge of the subject and were therefore more likely to know. Again all expert users and only 39% of content novice users felt they could distinguish the ready-to-be-learned links to follow. Surprisingly, 80% of content experts and 43% of content novice users felt it was easy to distinguish links they had visited. However, 80% of content experts and 60% of content novice users could distinguish the reference type links. This is a good result compared to distinguishing other types of links.

		Application		
		A	**B**	**C**
I could distinguish visited links	**Agree**	9	8	0
	Not sure	1	2	4
	Disagree	1	1	6
I could distinguish the ready-to-be-learned links	**Agree**	8	9	1
	Not sure	2	2	4
	Disagree	1	0	5
I could distinguish reference links	**Agree**	9	10	2
	Not sure	11	0	2
	Disagree	0	1	6

Table 8-7 Distinguishing Various Types of Links- Percentages

	Application	SD	Median	Mean	Min value	Max value
It was easy to distinguish links I ha already visited	**A**	2	0.81	2	1	4
	B	2	0.9	2.2	1	4
	C	4	0.8	4	3	5
It was easy to distinguish ready-to-be-learned links	**A**	2	0.7	1.9	1	3
	B	2	0.6	2	1	3
	C	4	1	3.8	2	5
It was easy to distinguish reference type links	**A**	2	0.5	2	1	3
	B	2	0.7	2.1	1	4
	C	4	1.1	3.8	2	5

Table 8-8 Distinguishing Various Types of Links- Descriptive Statistics

8.5.4.3 Ready-to-be-Learned (next-relevant) Links

Users from Applications A and B were asked if they followed the 'ready-to-be-learned' links because they were helpful and most users agreed. However, most also admitted to following the red 'not-ready-to-learn' links too.

8.5.4.4 Reference/ Generic Type Links

Out of the Application A and B user groups, even those who found the ANS confusing, felt that they could distinguish reference type links. These results are discussed further in the following sections.

8.6 Analysis of User Performance Logs

8.6.1 Introduction

This section investigates various aspects of the user performance logs. The logged data was used to investigate the differences in user performance between each application. The automatic logging software did not log the time variable, therefore this attribute was measured in seconds using a time-set function on the video recorder. Although this task took a great amount of time during the analysis stage it proved to be most advantageous. This was because during the analysis it was possible to establish the exact time that the user was looking at the screen and thus eliminate any time where the user may have left the desk or taken a break for any other reason.

The 'document reading' time includes the time users spent viewing and reading the documents, scrolling through the document, deciding which link to follow next and clicking on the link. The 'question reading' time refers to the amount of time users spent reading the tasks which were set on paper.

8.6.2 Following Generic Type Links

Application	Mean no. generic (type A, B and C) documents visited	SD	Max val.
A	6	5.6	21
B	5.8	3.7	11
C	2.8	1.9	6

Table 8-9 Mean Number of (Total) Generic Type Links Followed in Each Application

Application	Type A generic	SD	Max val.	Type B Generic	SD	Max val.	Type C generic	SD	Max val.
A	2.4	1.66	4	2.4	4.5	17	1.1	1.5	5
B	0.6	1	3	2.6	2	7	2.6	2.9	8
C	0.3	0.67	2	1.8	1.6	5	0.7	0.67	2

Table 8-10 Mean Number of Different Generic Type Links Followed in Each Application

Type A generic documents; Mean number of generic topic summary documents opened
Type B generic documents; Mean number of generic glossary documents opened
Type C generic documents; Mean number of generic image documents opened

Table 8-9 shows the average number of (total) generic links that were followed by each application user group. Table 8-10 shows the mean number of generic links followed by each application user group broken down by the three generic link types.

Both tables clearly indicate that the Application C user group followed a much lower number of generic links compared with users from Applications A and B, which have similar means. Again, the maximum number of generic links followed was only 6 for Application C users and up to 21 for Application A. This is not surprising as the Application C users could not distinguish these types of links. However, for

Application A and B users it was the opposite case as they could distinguish the generic links and therefore were more inclined to follow them.

The ANOVA test was carried out to establish if a difference existed between the total number of generic links followed by each application user group. The result of p=0.16 showed that the differences were not significant.

However, the ANOVA test (p=0.001), showed a significant different between the number of Type A links followed by each user group. Therefore, an independent t-test was carried out to establish where these differences lay. The only significant differences were found between the number of Type A generic links followed between Applications A and B (p=0.04) and Applications A and C (p=0.01). Therefore, we can say that the number of generic links followed by Application A users was slightly higher than for Applications B and C. However, a significant difference was also found between the number of Type A generic links followed between Applications B and C (p=0.04).

80% of content experts and 91.3% of content novice users followed 1 or more generic links (all types). 70% of content experts and 35% of content novices followed one or more Type A generic links. 70% of content experts and only 35% of content novice users followed one or more links to glossary terms on the current topic. 40% of content experts and 83% of content novice users followed one or more generic glossary type links (i.e. Type B).

60% of content experts and 60% of content novice users followed between 1 and 8 links to images. Table 8-11 shows the percentage of users who followed one or more of each type of generic link. The results show that the content expert users tended to follow more Type A and B generic links compared to novice users, whereas a similar number followed links to images and diagrams. The results found for Type A generic links could be due to the topic introduction pages having enough information on the topic for the current knowledge levels of the novice users. Experts however, may have needed to recap on existing knowledge.

	% of users who followed one or more generic links		
	Type A	Type B	Type C
Content Novice	35	35	60
Content Expert	70	70	60

Table 8-11 Mean Number of Generic Type Links Followed by Each SM Group

65% of content experts and only 30% of content novice users did not follow any Type B generic, glossary links. Experts may not have required the extra help or the novice users may have needed further basic information. 60% of novice users visited between 1 and 3 (and 22% between 4 and 17) Type B generic glossary links, whereas only 40% of experts followed between 1 and 3.

8.6.3 Time Spent Viewing Generic Type Documents

	Type A	SD	Max val.	Type B	SD	Max val.	Type C	SD	Max val.
App. A	12.4	11.9	45	4.7	6.5	21.6	3.7	4.7	16.5
App. B	5.3	9	22.3	4.5	3.5	10.3	4.5	4.3	11.5
App. C	1.4	3	8	5	3.9	14	5.2	6.9	20

Table 8-12 Mean Time Spent Viewing Generic Documents

Type A generic documents: Mean time spent viewing generic summary documents: these gave an overview/ summary of the topic.

Type B generic documents: mean time spent viewing generic glossary documents

Type C generic documents: mean time spent viewing image documents

Table 8-12 shows the average time, S.D. (Standard Deviation) and maximum time in seconds each application user group spent viewing each type of generic document. The table shows that the users of Application A spent a greater amount of time per topic summary document, i.e. Type A, at 12.4 seconds compared to Application C users at 1.4 seconds. However, the time for the Application A users does have much higher SD.

Interestingly a high positive correlation was found between the number of generic links that users followed and the average time they spent viewing Type A generic documents (r= 0.593). Also, a very high positive correlation was found between the mean time users spent viewing Type C generic documents and the total session time (r= 0.536). Again, the correlation found between the time users spent viewing Type C generic documents and the mean time users spent viewing documents (i.e. not including generic documents) was also quite a significant result i.e. p= 0.352.

Pearson's correlation coefficient p= -0.465 shows a high negative relationship which implies that Application C users spent more time viewing Type A generic documents than the other two groups.

Most novice users spent less than 13.3 seconds per Type A document and all expert users spent more time per Type A generic document. 50% of content experts spent between 4-5 seconds viewing images and 60% of novice users spent between 6-20 seconds viewing the images, perhaps because they need more time to decipher the images.

8.6.4 Number of Nodes Visited

Application	Mean no. of nodes visited (minus all generic documents)	SD	Min value	Max value
A	24	7.2	17	40
B	33	12.7	18	50
C	26	4.2	23	37

Table 8-13 Mean Number of Nodes Visited

As can be seen from Table 8-13, Application A users visited a fewer number of documents compared with users of Applications B and C. This was expected because they were not required to follow 'more complex information' type links within each document. More complex information type links in Applications B and C were links to complex information that was presented by shading out in Application A. The mean number of documents visited by Application B users was the highest at 33. This was expected as users were encouraged to follow the 'more complex information' type links and hence visit more documents. However, the mean number of documents visited by

Application C users is far less than was expected. This may be due to users of Application C not feeling safe navigating too far into the hypertext, which would ultimately result in them following too many links and possibly getting lost.

The ANOVA test was carried out to establish if the difference between the number of nodes visited (not including generic documents) visited between each application user group was significant, the result of p=0.06 shows that this result was significant at the 10% level.

Therefore, the independent t-test was carried out and the main significant difference was found between Applications A and B at p=0.04. (With equal variances not assumed). This reflects the nature of the applications i.e. users of Application B were expected to follow the 'more complex information' type links to documents containing complex information whereas this information was present within the document in Application A as shaded text.

Again the relationship between the total time users spent reading the questions and the average time users spent viewing a document (not including generic documents) was a highly significant positive correlation (r=0.712), i.e. the longer a user spent reading the questions the longer they spent viewing a document.

It was also found that quite a high percentage of content novice users, i.e. 51%, visited a maximum of 25 documents (not including the generic documents) or less. Whereas 80% of expert and 47% of content novice users visited a total of 25-50 documents, i.e. experts tended to visit more documents.

8.6.5 Time Spent Viewing Nodes

Application	**Mean time spent on each document opened**	**SD**	**Min value**	**Max value**
A	28	11.7	18.3	64
B	22	6.6	15.2	33.84
C	30.6	7.8	16.9	40.6

Table 8-14 Mean Time Spent Viewing Nodes

Table 8-14 shows the mean time spent viewing nodes by each application user group (it does not include generic type documents). It can be seen that the users of Application C spent the greatest amount of time per document at 30.6 seconds and had the lowest S.D. Application A users were a close second at 28 seconds and Application B users spent the lowest amount of time with 22 seconds. The S.D. for Application A was the highest, implying a large time range for which different users spent viewing documents while using Application A. The higher result for Application A was expected due to the nature of the adaptive technique i.e. the shaded text was added to each relevant document for this application; whereas for the other two applications more complex information was on a separate short document. This explains the lower mean time users spent viewing documents in Application B. The higher SD for Application A can be put down to the fact that some users did read the shaded text and some did not. The high time for Application C could be put down to the users of this application being given no assistance of any sort from the application, and therefore feeling the need to read each document in more depth.

The ANOVA test (p=0.01) showed a significant result. Therefore, we can say that there is a significant difference between the time users spent viewing documents within each application.

In general most content experts spent less than 18 seconds per document. However, the range was wider for the content novice user group. 70% of the content experts spent less than 23.9 seconds per document, whereas only 34% of novices spent less than 23.9 seconds per document. On the whole, most novice users spent more than 23 seconds per document i.e. more time than content experts.

Regarding the total time users spent viewing the questions, i.e. question reading time, all content experts spent less than 194 seconds in total and 47% of novice users spent 196-261 seconds reading the questions.

8.6.6 Re-visiting Nodes

Application	Mean No. of documents revisited	SD	Min val.	Max val.
A	4	3	1	11
B	6.6	5	1	16
C	5.3	1.5	3	7

Table 8-15 Re-visiting Nodes

Table 8-15 shows the mean number of documents re-visited within each application. Application A users re-visited the least number of documents compared with the other two groups. Application B had the highest mean but the mean was slightly less for Application C. However, Application C had the lowest SD. It is interesting to note that the minimum number of nodes re-visited by Application C users was 3 as opposed to 1 for each for the other two applications. Also, for Application C the highest number re-visited was 7 as opposed to the much higher values for the AH applications. This indicates that although Application C users did re-visit documents they were cautious, as the maximum value of 7 indicates, and the users of the AH applications navigated more freely.

Even though it was a subjective opinion from the users it was interesting to find that all Application C and 10 Application B users felt that they needed to re-visit learning material. However, only 8 users from Application A felt they needed to re-visit nodes. The ANOVA test produced a result of $p=0.22$ which is not significant, i.e. the number of documents re-visited by each user group was not significant.

8.6.7 Following Not-Ready-to-be-Learned Links

Table 8-16 shows the difference between the number of not-ready-to-be-learned links followed by Applications A and B users.

Application	Mean no of not-ready-to-be-learned links followed	SD
A	2	1.9
B	2.33	1.32

Table 8-16 Following Not-Ready-to-be-Learned Links

The table clearly shows that there is not a great deal of difference between the number of not-ready-to-be-learned hyperlinks followed by users from either application. 3 users from Application A and 2 from Application B did not follow any not-ready-to-be-learned links. 3 users of Application A and 1 from Application B followed one not-ready-to-be-learned link. 3 Application A users and 2 Application B users followed 4 or more links. The questionnaires revealed that the users who followed not-ready-to-be-learned links did so because they felt that the answer was at that destination document. On the whole most of these users were right. This is explained in more detail below.

It was also found that 91.7% of the novice user group and only 66.7% of the expert user group followed one or more not-ready-to-be-learnt links, i.e. the content novice users followed more not-ready-to-be-learnt links than content experts.

The independent t-test (p=0.65) showed that the difference between the number of not-ready-to-be-learnt links followed by Application A and B user groups was not significant.

8.6.8 Following 'more complex information' Type Links

Application	Mean no. of 'More-information' links followed	Min val.	Max val.
B	3.8	0	16
C	2.9	0	7

Table 8-17 Following 'More Complex Information' Type Links

Table 8-17 only refers to Applications B and C as the documents were designed to encourage users to follow 'more complex information' type links. The table shows that users of Application B followed a higher number of 'more complex information' type links (16) than users of Application C (7). Showing again that Application C users were more cautious- maybe because they could not distinguish these links or did not find them important. Again, the independent t-test (P=0.5) showed that the difference between the number of 'more complex information' type links followed by Application B and C user groups was not significant.

Interesting, only 30% of content experts and 57% of content novice users followed 'more information' type links.

8.6.9 The Mean Number of Tasks Answered Correctly

Application	Mean no. of tasks answered correctly	SD	Mean no. of tasks answered incorrectly	SD
A	13	1.4	0.6	0.7
B	12.2	1	0.2	0.4
C	12.9	0.87	0.5	0.7

Table 8-18 Mean Number of Tasks Answered Correctly

Table 8-18 shows that there was a slight difference between the mean number of tasks answered correctly between each application group. However, this difference was not found to be significant.

8.6.10 User's Anticipation of Destination Documents

Two questions on the task list were phrased in such a way that users could make a relatively accurate guess as to which link they could follow to find the answer. One was a question on Radiocarbon Dating that contained the same keyword as a not-ready-to-be-learned hyperlink. Users were expected to ignore the advice provided by the system and follow the not-ready-to-be-learned link. 5 Application A users, 7 Application B users and 5 Application C users went directly to the right document. This shows that even users who follow the advice provided by the application will override the advice if they are sure they can find the right document using a more direct path. A second question of this nature was on (Obsidian) Hydration Layer Measurement. However, this question was phrased less obviously and the results were quite different. Only 1 user from Application A and 1 from Application B felt they could reach the answer directly from the first page. However, 5 Application C users attempted to reach the documents directly. This maybe because they were not restricted by the adaptation and were freer to follow any link. On the whole the topic on Obsidian posed a few more problems than the other topics. Most users spent the greatest amount of time viewing each node and visited more nodes within this topic.

8.6.11 The Usability of the Shading Method of Adaptive Hypermedia

An important aim of the study was to investigate the usability, effectiveness and acceptance of the shaded method of AH used within Application A. The video recording was used to ascertain if users were viewing i.e. reading or scanning through the shaded text. This was also a more direct way to determine if the shaded text was intrusive in any way to the reader. On the whole, a high percentage of users of all three applications tended to run the mouse over the lines they were reading on the document. Therefore this made it easier to determine if users were reading or at least viewing the shaded text. The post-session questionnaire results showed that the users of Application A were extremely positive in accepting this new form of AH. Watching their interaction with the application on the video film reflected this. The video recordings showed that a very high percentage of the content novice group viewed the shaded text within each topic, even though they were told they were not required to. (By the same token it was interesting to find that some content experts did not view the shaded text within certain topics). All Application A users, regardless of their knowledge level read or viewed the shaded text within one or more document. That is, all users viewed the shaded text. 75% of content novice users and only 40% of content experts felt that the shading technique was useful.

8.7 Discussion

8.7.1 AH in Education

In general, computers and learning software seem to be accepted by students as a possible aid to learning. This was reflected in that all the participants in the study said they enjoyed using a computer while studying. It was encouraging to find that all the participants felt AH was a beneficial concept and could compliment conventional lectures. It was also encouraging to find that an increased number of participants felt that they would use a CAL application after the evaluation session. This showed that the users felt comfortable with the application they used and that AH was a more beneficial concept than they may have originally thought. When asked if tutors should still be present while students used CAL applications the percentage that agreed fell after the session. This shows that the users formed a positive attitude once they had used the CAL applications and implies that the applications provided significant guidance that was acceptable by the users.

In the pre-session questionnaire users were asked if they preferred using books as they could access any chapter or section. A very high percentage agreed, implying that they do not accept the idea of hiding information when studying, even if they are not knowledgeable about the subject. This result was reflected in the answers to, 'should information be made available if you are not yet ready for it?' where a very high percentage agreed. On the other hand, a very low percentage of users felt that lectures were useful because they only exposed them to limited information. These results highlight the importance of making all material available within a learning application and could explain the wide acceptance of the shading method of AH. This may also be why users have not accepted hidden links in previous studies.

8.7.2 Comparing Applications

It was very interesting to find that although Application A users had a higher overhead compared to the other two applications, a higher percentage of Application A users agreed that they could find the information they required quickly and easily. Even though they had an increased overhead with the ANS and shading. These results indicate that the presentation of information using shading does not lead to confusion or result in information overload as users still felt that they could find relevant information.

On the whole, most users found the shaded fragments of text useful. This was reflected in the user's opinions, i.e. that shading did not disrupt their flow of reading or obscure the text in any way. It was not expected that a higher percentage of content novice users found shading text useful compared to content experts. This implies that shading could be a very useful technique to encapsulate complex material that is exposed to novice users as opposed to expert users.

It was encouraging to find that a higher percentage of novice users (at least) viewed the shaded text in Application A compared to the number of content novice users who followed the 'more complex information' type links in Applications B and C. This indicates that links could still prohibit users viewing extra material. Although it was difficult to assess if users actually read the shaded text or not, the video recordings did show that most users of Application A did at least skim over the shaded text. This could be seen as the users (in fact, most of the subject who took part in the study) tended to run the cursor over the sections of text they were reading. In some cases the shaded text was on the screen as users were looking at the screen, again implying that the shaded text was at least viewed/ skimmed over even if not read.

9 users of Application A felt that the guidance provided by the their application was useful with only 5 Application B users and 2 Application C users agreeing. The result for Application C was not surprising, as the users had no support or guidance provided by the application. However the result obtained for Application B users was surprising as they were provided with ANS and had no other overheads. This implies that users accepted that a combination of AH techniques is required for a CAL application to be beneficial and that ANS alone is not enough and that the shaded text and ANS was a supportive combination. It also implies that even a well-designed hypermedia application does not provide relevant support.

A higher number of Application A and B users felt they had control over their applications, whereas a lower number felt this way regarding Application C. This was a positive result indicating that AH provides users with some control and guidance compared to hypermedia applications. This highlights the importance of providing users with a sense of control which ultimately affects their attitude, confidence and the way they manipulate and navigate around the application.

On the whole, a high percentage of users said they did not find their particular application confusing to use and they did not feel lost or disoriented. This implies that ANS was not confusing to users (even for those who had used AH for the first time). This is an encouraging result. However, it was found that a slightly higher percentage of content novice users felt that ANS was confusing compared with content experts. This problem arises probably because content novice users are trying to learn new material and at the same time attempting to understand the ANS coding. One solution to this issue is to design simpler annotation to make it easier for content novice users.

On the whole, a higher percentage of Application A user's found the ANS useful. This maybe because they felt that it assisted the shading technique, which alone may not be as useful. It was positive to find that even users who found the ANS confusing to use felt that the guidance provided by their application was useful. Again, this implies that although ANS is useful it requires refinement. It is important to point out that although some users did feel that the ANS was confusing most could still distinguish the different types of links.

With respect to generic links, a higher percentage of content experts followed generic links compared to novice users. It was not surprising to find that Application C users followed a fewer number of generic links compared with the other two application groups, probably because they could not easily distinguish between the different types of links. Interestingly, twice the percentage of content experts compared to content novice users said they could distinguish between the different types of links.

8.7.3 SM Variables

This study also investigated some variables that could be used by the SaDy SM to predict user expertise or assist in guidance decisions based on performance indicators. The study emphasised the relevance of users educational background in a SM i.e. the subject of study. Another variable that was found to be relevant for a SM was the number of nodes visited. Again, this study like the first empirical study also highlighted the importance and possible relevance of 'time' as one type of performance indicator for a SM in an educational AH system. These differences between the domain expert and domain novice users were highlighted in the performance logs. It was found that a higher percentage of content experts (compared with novice users) spent a greater amount of time viewing the topic summary documents. This maybe because experts still needed to recap on some of the basics. This type of result is important as it indicates the main differences between content expert and novice users, and the type of information they are more likely to require from a learning application. This is also relevant in the case of expert users with respect to their retention of knowledge of the concept or topic as it shows that even content experts require access to basic information etc. This could also be relevant depending on when they last studied the subject.

Application A users on average spent a greater amount of time viewing documents, but that was expected due to the nature of the application i.e. the shaded text. This implies that the users read the shaded text and is supported by the evidence from the video analysis. Application C users spent the greatest amount of time viewing documents. Again, this was probably due to the lack of guidance provided by the application. Also, users of Application A tended to spend a greater amount of time viewing Type A generic documents, i.e. topic summary documents and Application C users spent the greatest amount of time viewing Type C generic documents i.e. images.

On average, content novice users spent less time viewing documents compared to the content expert user group. This could be because they were unfamiliar with the content and needed to read it thoroughly.

Application A users visited the least number of documents compared with Applications B and C. Again, this was expected. Application C users visited far less documents than expected. On the other hand, Application B users visited the highest number of documents. Implying that they did not feel apprehensive about navigating around the application, again, probably due to the additional guidance provided by the ANS. Application B users also followed a higher number of 'more information type' links compared to Application C users. Again, probably because they found it easier to navigate around the application due to the ANS. Also, a fewer number of content experts followed 'more information' type links compared to novice users. This maybe because they already knew the material. An increased number of content novice users followed not-ready-to-be-learned links compared with experts. A fewer number of Application A users attempted to follow the not-ready-to-be-learned links compared to the users of the other two applications. This could be because they had the shaded text to consider and therefore followed the adaptive advice provided by the system more closely. Again, it was encouraging to find that Application A users re-visited fewer documents compared to users of Applications B and C. Although this measurement is not an ideal indicator within this study due to the nature of Application A.

Although a comparison of learning using the various applications was not an aim of the study, a basic analysis found that no significant difference was found between the number of tasks that each application user group answered correctly.

8.7.4 Windows Management

It is important to note here that all the users in the trials had some level of experience using a computer and therefore all had good knowledge of windows management. Therefore it was not surprising to find that all users changed the size and the position of the windows to suit themselves. Most of the users arranged the windows in the middle of the screen. One user preferred narrowing the windows to read the documents due to eye sight problems. This user felt that it was important to be able to change the font size i.e. to make the interface adaptable as well.

8.7.5 Identifying Various Navigational User Types in AH

By thoroughly analysing the patterns of user trails through each application, it was possible to extract three distinct types of navigational patterns, which could be used by the SaDy SM to provide guidance to users before they get too confused or irritated;

- Type 1: Followers
- Type 2: Repeaters
- Type 3: Roamers

Type 1, 'the followers', can be divided into two sub categories;

1. The right starters; users who followed the advice most of the time.
2. The wrong starters; users who first attempted to ignore the advice given and follow the not-ready-to-be-learned links. Then realised that they were getting confused or lost and decided to follow the advice.

For users of Applications A and B, some users first followed the not-ready-to-be-learned links (but only up to a maximum of three in a row) and then decided to return to the introduction page of that topic and from that point, follow the advice given by the system.

Type 2, 'the repeaters', are basically the reverse of Type 1 users. These are the users who followed the advice given by the system and if necessary repeated the process, these users trusted the advice provided by the system.

Type 3, 'the roamers' are again divided into two groups;

1. Users who followed the advice given by the system until they got frustrated at which point they decided to randomly access any document.
2. Users who did not follow the advice provided by the system (in the form of ANS, shading or the table of contents) and randomly accessed the documents throughout the session.

This type of user was more prominent within the Application C user group. These users first tried to go straight to the document, which they felt contained the answer.

When they could not find the relevant information they attempted to randomly follow links and open documents. 54% of Application C users followed this pattern of browsing. It was found that from the Application A and B user groups, the users who did not follow the system's advice resulted in getting lost and visiting more documents. No user of any application attempted to use the search facility on Microcosm. This maybe because all users were told at the start of the session to use the highlighted links to navigate around the hypertext.

8.8 Conclusion

Most of results have a higher SD for the AH applications and lower for the Control Application. This implies that the users of Application C worked in a more predictable and restricted way compared with the AH user groups. They were not as adventurous maybe because they had no guidance and were scared of getting lost, i.e. the increased cognitive overhead that comes with having to remember where you were, need to go or have already been.

The shading method of adapted hypermedia was positively accepted by most of the users regardless of their content knowledge. However, content novice users found it more useful than experts did. This implies that it is more useful to shade material that is not currently required by the user or is not at their expertise level. This could be complex material for novice users and basic material for expert users. The shading of the fragments of text did not appear to disrupt the flow of reading or add to the cognitive overhead in any other way.

ANS does add an extra overhead to users in the form of remembering the coding i.e. what colours, fonts etc mean. However, ANS reduces overhead when users need to know which links to follow next or which links they have visited etc.

The study also highlighted some important variables for the SaDy SM or any SM. The investigation showed a significant difference between the performance of content experts and content novice users. Implying that variables such as the following could be used in the SaDy SM, type and number of generic links followed, number of nodes visited and time spent on nodes. Another important issue to come out of the study was the significance of the 'time' variable in the SaDy SM, i.e. the time a user is viewing a relevant document, generic documents etc. as this can provided the system with some indication of the user's current experience and performance levels.

The next chapter will draw the thesis to a close by presenting the conclusions of the research. It will also summarise the contribution of the research to the field of AH and finally, discuss some ideas for future work.

Chapter 9 Conclusions and Future Research

9.1 Conclusion

AH is a relatively new and dynamic area of research. This thesis has attempted to present and describe in detail a new adaptive presentation technique in AH. The study revealed and highlighted some interesting and exciting issues. A summary of the conclusions is documented below.

1. The thesis introduced a new adaptive presentation method of AH i.e. shading fragments of text, a method that is currently used in conventional textbooks.

2. The shading method of AH allows users of varying content and computing skill levels to use the same application and possibly view the same material i.e. no information is hidden from the user.

3. The empirical experiments carried out in the study used a variety of conventional interface evaluation techniques. For example pre/ post-session questionnaires, interviews, observation and software logging. The sessions were carried out using a formal as approach as possible given the environmental and time constraints.

4. The evaluation sessions exposed most of the subjects to AH for the first time. As a result the acceptance of adaptive technology in an educational environment increased as well as the general awareness of AH techniques. This was strongly highlighted in the comparison between the pre and post-session questionnaires, where user attitudes shifted towards the acceptance of AH after only one session using AH.

5. The study highlighted that hypermedia within the educational paradigm is made more effective if adaptive technology is incorporated. However, users felt that ANS alone was not beneficial and preferred combining ANS with a complimentary adaptive presentation technique such as shading. Implying that ANS and shading are complementary AH techniques. The study also showed that as well as shading being an effective AH technique the users appeared to enthusiastically accept it.

6. Another advantage of the shading method is that it did not appear to interfere or disrupt the user's flow of reading the on-screen text.

7. A specific advantage of the shading method of AH is that it reduces the number of highlighted links on the document resulting in reduced cognitive load on the user.

8. Users of the AH applications (especially the ANS and Shading Application) visited fewer documents, re-visited fewer documents and followed more generic links.

9. The experiments also revealed a distinct difference in the performance of users working on adaptive hypermedia applications, which was found to be related to the user's current content knowledge levels. A distinct difference was found between the way that content experts and content novice users worked with the AH applications. On average, content experts visited more documents and spent less time viewing documents compared to content novice users. Content expert users were also more likely to follow not-ready-to-be-learned links than content novice users.

10. Although SMs are a complex area of AH systems, this study investigated some basic SM issues and as a result attempted to construct a basic framework for a SM for use within the educational paradigm, i.e. the SaDy SM.

11. The study also began to investigate the notion of time as a variable in the SaDy SM. The results from both empirical studies demonstrated that time could be used as a performance indicator. However, more detailed work needs to be done in this area.

12. Although the acceptance of AH was overwhelming, it is clear that as far as educational hypermedia goes we need to design systems with more intelligence and guidance abilities. However, it is important to ensure that the user feels in control and not overwhelmed by the guidance provided by the system. This

need was reflected in the difference in performance between the users of the AH applications and the control Hypermedia Application.

9.2 Summary of Contribution

This research introduced a new method of AH- shading fragments of text- which may be classed as an adaptive presentation technique. This novel approach to AH was designed, developed and implemented within the research. Two empirical studies were designed and carried out to test the shaded method of AH within an educational environment. During the course of the research, the SaDy SM framework was developed.

The empirical evaluation has shown that AH and, more specifically the shading method of AH can be used effectively within an educational environment. In addition it was encouraging to find that this method was well accepted by the users. Although some of the measurements are a result of the nature of the application itself, the fact that these differences exist and are worth mentioning for example, while using the shaded method of AH the users were able to find the relevant information easily, they visited fewer documents, they re-visited fewer documents and they got lost less often. They were more likely to view material on the screen than follow a link to it. The most positive feedback was from the users themselves, who said they preferred the Shaded Application to the ANS and control Hypermedia Applications. Although the shading method of AH was accepted by the users and found to be useful, it does require refinement and further work, this is discussed in the next section.

9.3 Future Work

9.3.1 Shading Vs ANS on the WWW

The next step in this area would be to experiment with dynamic shading and ANS in the WWW environment. The WWW is a distributed environment so users can access applications from anywhere on the network. There is a need to develop methods of establishing the user profile on-line, and for the system to present the adapted application (shading or ANS) automatically as a result. The DLS (Distributed Link service) (Carr et al., 1998) enables links to be stored in separate linkbases and be applied to WWW documents (via a proxy). Work has also been carried out to investigate the use of an agent based framework where a agent provides the user profile to the DLS to determine which linkbase to use (Bailey, 2000). Any future work in this area would also need to consider how to dynamically shade fragments of text in a HTML (hypertext markup language) documents maybe using XML (extended markup language).

9.3.2 Dynamic Shading of Fragments of Text

Future work using the shading method of AH could be geared towards investigating dynamic shading of fragments of text for different purposes. For example as the user progresses through the information base, basic learnt concepts could be shaded out. Another important issue related to this area is to investigate what could be considered as a suitable minimum and maximum length of the shaded text fragments. As well as the maximum and minimum number of shaded fragments of text within one document; Can a document contain too many shaded fragments? Does it depend on the type or level of information that is shaded, i.e. learnt, complex, basic etc? This leads into the next point, which is that studies also need to be carried out to establish if it is more beneficial to shade complex text, basic text, learnt text or only complex calculations or formulae. As this study found, these results would probably depend on the subject's content knowledge.

9.3.3 Shading Fragments of Text

For an application supporting static shading, it is the subject expert who decides what material would be considered complex for a user group ie domain expert or novice users, depending on current knowledge levels. This information would result in particular sections of text being shaded during the design of the application. For example, an application for first year undergraduates would consider calculating the radiocarbon date for a particular artefact complex for novice users. Therefore all information relating to this calculation would be included in the document but shaded in grey. When a novice user comes across this document during the course of the interactive session, they would see the calculation information with the background shaded and would know that this was irrelevant information for their current knowledge level.

To implement an adaptive shading mechanism, the UM model would need to be dynamic. An on-line pre-session questionnaire should be used to obtain user information on current knowledge levels etc. All the material can be split into fragments of text therefore, the sections of text would be treated as in conditional text which is either hidden from the user or is displayed to them depending on the users UM etc. The domain model holds the relationships between all the documents and or the fragments of text. Dynamic shading could be represented in one of two ways. Firstly, by dynamically shading a fragment of text before it is presented to a user, or by holding two versions of each fragment of text (that makes up a document) and either the shaded version of the text would be presented to the user or the unshaded again, depending on the current state of the users UM and skill level.

Therefore if the user requires complex information to be hidden as they progress through the application and pick up more knowledge, the shaded text that is presented to the user will be different depending on how much knowledge they have picked up during the course of the session or on which documents they have already visited and therefore the application has decided they have new knowledge in this area. The UM updated during the session and the UM knows what knowledge the user has picked up either just by visiting the document or by tests. Therefore before displaying the fragments that make up a document the system accesses the knowledge from the UM and decides what should be shaded and what should not be shade. It works exactly the same as conditional text.

As the user goes through the documents or passes tests on a topic, the system assumes that the user has learnt enough about a particular concept or topic to be ready to be exposed to slightly more complex material, this complex fragments of text would be presented without shading. If However, a user reaches the complex material without passing a test or visiting the relevant pre-requisite documents, the system will assume that they have not learnt and will display the fragments of text or document with the complex text still shaded.

Future work could also look into developing a design methodology for dynamically shading fragments of text. This methodology would need to rely heavily on the domain model and the user model.

9.3.4 Allowing Users to Make Decisions

Future studies could consider giving users the freedom to make adaptation decisions. For example, allowing the user to decide the colour of the hyperlinks, the maximum number of the hyperlinks to be displayed, the type of font used to display links or allow a combination to be used. The user could even be given the choice of which adaptation technique to use, e.g. hidden links, ANS, link sorting etc. It would be interesting to investigate if users can be trusted to make correct decisions. The results could then be used to investigate the development of SMs and the value of allowing users to access and override dynamic SM decisions based on their progress.

9.3.5 Acquisition of Knowledge

An important issue in AH and more specifically in student modelling, is the acquisition of knowledge while using the AH application and the retention of user's knowledge. One way to investigate the impact of the retention of user's knowledge is to carry out a follow up session a few weeks or months after the original evaluation session. However, this maybe difficult to carry out due to time and resource constraints.

For AH systems, explicit user feedback can play an important role in verifying the users SM or to assess the acquisition of knowledge; this method was used by Kaplan et al. (1993). However, asking the users to provide feedback can be useful only if they are willing to give the time to answer correctly.

9.3.6 Student Models

This research introduced the SaDy SM framework for use in educational AH systems. An advantage of the SaDy SM is that it has a simple and modular structure that makes it scalable and therefore useable for different types and sizes of applications. It would therefore be interesting to use this model in future experimentation. The variables in the SaDy SM could be used to explicitly gain current information about the user at the start of a session. This information is used to dynamically allocate users into a generic SM group. As the user progresses through the session the dynamic SM can be updated and used to make adaptation decisions. In effect, the SM will learn about the user by updating the dynamic variables during the session resulting in a more individual SaDy SM.

In future work it would also be useful to investigate the usability of various combined AH techniques in education. Where one AH system would be designed to use the SaDy SM and the second system would not use any dynamic SM; this system would act as the control application. It would also be interesting to compare the SaDy SM with other similar SMs, i.e. basic SM for use in educational applications.

Future studies into student modelling should also be aimed at investigating the use of time as a variable of the SM. The studies presented in this thesis found that time could be used as a performance or knowledge indicator. However, further studies may require an in-depth investigation over time into user logs, e.g. by comparing different levels of users and their navigational patterns. This type of work may also require some specialist equipment such as an eye tracker that could be used to inform the SM when the user is viewing the screen.

9.3.7 ANS – Hidden Links

As users are introduced to hypermedia and then to AH systems, hidden links may become more acceptable. Although hidden links can be vital in reducing cognitive overload and reducing the visual impact of the screen, users have not readily accepted this method. Therefore, the issue of hidden links in educational AH systems needs to be approached in a more novel way. Microcosm is an excellent system to carry out such experimentation because all generic (glossary type) links in Microcosm are hidden by default. This can also be achieved using the DLS. One approach is to present these hidden links in italic font and no highlighting, so that users can distinguish these links. This will result in a screen, which contains fewer highlighted links and where the hidden links are presented in black italic text, which can only be recognised as links by their font style.

9.3.8 AH and User Authored Links

An interesting progression in AH research would be to consider the implications of user-authored links. For example, Microcosm allows a user to create their personal set of links in addition to the original authored links in the application. Therefore, the system is left with an application that has at least two sets of links; one set created by the author and another set by the individual user. How would an adaptive system handle or model this extra input? The SM could either vet the new link and respond by classing it as a valid relevant link or class it as an invalid irrelevant link depending on the domain model. Once this has been decided the system would then need to add this information to the SM and the domain model. Alternatively the system could dynamically create a new, separate domain model. What would be the implications for any adaptation decisions?

9.3.9 Adaptive Navigational Aids

Another important point to consider with AH systems is the use of navigational aids such as visual maps or a table of contents that can assist users in navigation and maybe even more importantly, help users to build correct mental models of the hypertext structure. These maps should be designed to be adaptive and interactive, to reflect the user's current goals and actions.

9.4 Concluding Remarks

This programme of research has developed a new and novel technique in AH, i.e. shading fragments of text, with an emphasis on educational hypermedia applications. The study also presented the SaDy SM, which is only a basic SM framework.

Before using the AH applications, users were wary of such technology but once they had used it, they accepted that it has advantages. This initial attitude will eventually become more positive with the increased use of technologies such as the WWW etc., which is helping users to accept basic hypermedia technology. In summary, as the need for managing and navigating through information in an educational environment increases, the awareness and the acceptance of AH technology will also increase. While there is still a great deal of research to be undertaken and questions to be answered in the field of AH, this research has endeavoured to explore some of the issues facing the design and ultimately the acceptance of hypermedia in education.

Appendix A Initial Survey Questionnaire

This questionnaire forms part of my research into the evaluation of the Microcosm interface. Thank you for your time.
If you have any queries please contact Mrs. J. Hothi on jh95r@ecs
(Sections 1-8 are designed to establish a connection between the skill of a computer user and their work/ home environment)

1. Sex (please circle) Male Female

2. Age (please circle) 21 or under 22-25 26-30 31-40 41-50 over 50

3. Country in which you went to secondary/ high school ....

4. Are you: (please circle)

Head of Department Academic staff Research assistant/ student

Undergraduate student Postgraduate student

5. Computer experience (please circle one number on each line)

a.

	expert	adequate	complete novice	never used
Microcosm	1	2	3	4
word processing	1	2	3	4
games	1	2	3	4
graphics	1	2	3	4
computer based learning	1	2	3	4
programming	1	2	3	4
Microsoft Office	1	2	3	4
e-mail	1	2	3	4
World Wide Web	1	2	3	4
authoring on the World Wide Web	1	2	3	4
authoring on the Microcosm	1	2	3	4
Windows management	1	2	3	4
typing speed	1	2	3	4
when using a computer I am: enthusiastic	1	2	3	4

b. Which input method do you prefer to use (please circle)

Keyboard Mouse

6. Personal Characteristics

a. Please circle the number which best describes you on the scale.

shy	1	2	3	4	5	extrovert
quiet	1	2	3	4	5	loud
pessimist	1	2	3	4	5	optimist
timid	1	2	3	4	5	overpowering
confident	1	2	3	4	5	unconfident
like working alone	1	2	3	4	5	in groups
I am confident using computers	1	2	3	4	5	unconfident

b. Please state your interests
....

c. Please state your habits (good or bad) when you are working on a computer (e.g. Never closing down windows which are not required)

....

d. What sports do you like playing

....

7. Home Environment

a. Do you live: (please circle)

Alone With one person

With less than 4 With 5 or more

b. Are there children below the age of 7 at home, if so how many.................

c. How much time do you spend watching television (hours).........................

d. How much time do you spend working on a computer at home (hours)......................

d. How much time do you spend on your hobbies or other leisure activities (hours).....................

8. Work Environment

a. What is the set-up of your workspace: (please circle)

Use a lab Office to yourself Share with 1 other person

Share with 4 or less people Share with 5 or more

b. Do you get easily distracted when you work on a computer

Yes No

c. Do you work with many (more than 3) windows open on your computer screen at one time

Yes No

d. Is your desk usually a mess

Yes No

9. Microcosm Use (please circle)

a. Have you ever used Microcosm Yes No

b. If you answered 'Yes' to question 9a.go on with question 9 if 'No' goto question 11

c. What do you use Microcosm for and which subject:

Authoring ....

Teaching/ tutorials ....

Learning ....

Browsing information ....

Reference information ....

Other (please specify)...

d. If you arrive at a document you do not require- do you know how to get back to the previous document

Yes No

e. If you answered 'No' to question 9d. would this be an important function to have available

Yes No

f. Do you enter Microcosm via: (please tick as many as are relevant)

Select a document window

Front end designed specifically for the application you are using

Other (please specify).....................................

g. Which of the following Microcosm functions do you use: (please tick as many as are relevant)

Highlighted (buttons) text within the documents

Compute links

Follow links

Show links

History

Search

h. Do you feel that Microcosm window management allows too much freedom:

Yes No

i. If you answered 'Yes' to question 9h.- What changes do you feel would help

....

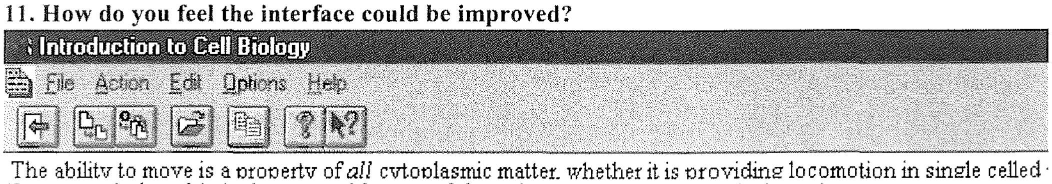
j. What functionality do you think Microcosm lacks for you to complete your task (be it anything, learning, authoring etc.) easily and quickly

....

10. Microcosm- a Presentation System (please circle one number for each question)

	strongly agree	**agree**	**neutral**	**disagree**	**strongly disagree**
a. Colours used for the interface					
too many colours	1	2	3	4	5
good contrast between colours	1	2	3	4	5
b. Layout of the interface is clear	1	2	3	4	5
c. Too many windows are on display at one time	1	2	3	4	5
d. Functions available on the interface are easy to use	1	2	3	4	5
e. I can find relevant information quickly and easily	1	2	3	4	5
f. I can feel lost and confused using Microcosm	1	2	3	4	5
g. The interface is easy to use by users of varying computer skill	1	2	3	4	5
h. Too much information needs to be memorised	1	2	3	4	5
i. Overall reactions to the interface					
interesting	1	2	3	4	5
stimulating	1	2	3	4	5
easy to use	1	2	3	4	5
j. Problems are easily solved	1	2	3	4	5
k. Messages given by the interface					
do clarify the problem	1	2	3	4	5
do indicate action to be taken	1	2	3	4	5
are relevant to my problem	1	2	3	4	5
l. Undoing operations is easy to do	1	2	3	4	5
m. Computer Aided Learning applications are effective learning tools	1	2	3	4	5

11. How do you feel the interface could be improved?



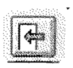
(Just a reminder- this is the general layout of the Microcosm using a RTF viewer)

a. How do you feel this could be improved

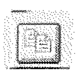
....

....

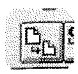
b. What do you think the following icons represent:

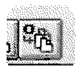


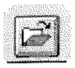
....

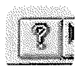


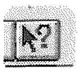
....


....


....


....


....


....

12. Intelligent Interface Issues- Please answer the following questions on the basis that the interface was intelligent and adapted to individual user needs.

a. When using a computer based learning application, more intelligent support is required from the interface in the form of:

Hints (Brief description e.g. Letting the user know they have missed something vital which is required before attempting a certain task) ☐

Guidance (Information giving a bit more help e.g. Letting the user know where they should look for that information) ☐

Help (More detailed help (e.g. Let the user know exactly what they need to complete a task) ☐

b. If any of the following input methods were available which would you prefer to use (please circle as many as are relevant)

Keyboard Mouse Speech recognition

Touchscreen Lightpen

c. When errors are made, I would like the following displayed? (Please tick 1 box)

Detailed information- Error message, cause and options available to rectify it	
Some information- Error message and the ability to retype	
No information- Just the ability to retype once realised error has occurred	

d. Assistance from the interface is most useful when ? (Please tick as many boxes as are relevant)

I ask application for help when I need it	
The application offer help when it feels I need it	
The application changes dialogue according to my computing skill level	
The application suggests documents I may need but have not yet read	

e. Which of the following status information would you find useful for a student? (Please tick as many boxes as are relevant)

What skill level you have reached on the running program	
The number of errors made	
Time taken to complete a task	
The windows you already have open	
Time spent reading documents	
Number of relevant documents not viewed	

Thank you for your co-operation

Appendix B Pilot Study- Pre-Session Questionnaire

Thankyou for your co-operation. Please answer the pre-session questionnaire. Then use the application to answer the set tasks. Once completed please fill out the post-session questionnaire. Please talk aloud and make any comments; good or bad that may help to improve the evaluation process.

Pre-session questionnaire

Please circle the relevant answers

1. Sex M F
2. Are you a confident person Y N
3. Are you a quiet person Y N
4. Are you a shy person Y N
5. Are you a pessimistic person Y N
6. Do you watch > 15hrs of television per week Y N
7. Do you spend > 15hrs on a computer per week (work and at home) Y N
8. What is your skill level on windows management novice expert
9. What is your skill level on the following applications
   - word processor novice expert
   - e-mail novice expert
   - web novice expert
   - MS office novice expert
10. What is your skill level on the application content i.e. radio carbon dating in Archaeology.
    novice expert
11. What is the average no. of windows you work with on screen at one time ....
12. How much help to assist with your task would you like access to from an educational application
    none some all
13. How would you like access to help (please tick one option)
    - ask the application when you feel you need it ....
    - the application offers when it feels you need it........
    - both the above........
    - none of the above.......
14. Would you prefer adaptation in hypermedia in the form of (please tick as many as you like)
    a. links to references within the document........
    b. only links to relevant task material within the document.......
    c. display content relevant to your skill level on the task with other parts of the document shaded out......
    e. allow access to all information available with all links possible......

Appendix C Pilot Study- Tasks and Post-Session Questionnaire

Using the Adapted Navigation Application
Tasks

There are 2 short tasks to complete the first using the Adapted Navigation Application and the second using the Adapted Presentation Application.

1. First, if you are a novice on radiocarbon dating log onto Microcosm using the user name 'link_novice' and password 'link_novice'.
3. Double click on pilot archaeology
4. Bring up the document management structure and click on 'content_novice' and then click on links.
5. Open the first document titled 'radcar1'. Now please complete the following task;

Please give as detailed answers as possible.
Double click on the highlighted text to activate links to navigate.
How are samples collected for radiocarbon dating?

What problems are encountered when using radiocarbon dating?

Post-Session Questionnaire

For the adapting of links application please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too few links			
there were too many links to definitions			
did you need to revisit documents			
presentation of the documents was clear			
this was a useful form of adaptation			

Using the Adapted Content Application

Tasks

1. Exit out of Microcosm and then Log on again, using the user name 'contNE' and password 'contne'.
2. Double click on pilot archaeology
3. Bring up the document management structure and click on 'contnet_novice' and then click on 'content'.
4. Open the first document titled 'radcont1'. Now please complete the following task;

If you are an expert on dating techniques in archaeology use all the text including the shaded parts in the documents to answer the question. If you area novice do not read the shaded parts of text.

Please give as detailed answers as possible.
What is carbon?

How accurate is radio carbon dating?

Post-Session Questionnaire

For the Adapted Navigation Application - please answer the following questions after completing the tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too few links			
there were too many links to definitions			
did you need to revisit documents			
the documents contained too much text			
this was a useful form of adaptation			
shaded text should be available to novice users			

Appendix D First Empirical Study- Pre-Session Questionnaire

This is an evaluation that will be looking into the use of adaptive hypermedia in education. All users must answer this questionnaire
Thankyou for your co-operation.

Pre-session questionnaire

Please circle the relevant answers

1. Sex M F
2. Are you a confident person Y N
3. Are you a quiet person Y N
4. Are you a shy person Y N
5. Are you a pessimistic person Y N
6. Do you watch > 15hrs of television a week Y N
7. Do you spend > 15hrs on a computer a week Y N
8. What is your skill level on windows management novice expert
9. What is your skill level on the following applications

word processor	novice	expert
e-mail	novice	expert
web	novice	expert
MS office	novice	expert

10. What is your knowledge level on dating techniques in Archaeology
novice expert
11. What is the average no. of windows you work with on screen at one time ....
12. How much assistance would you like access to, from an educational application
none some all
13. How would you like access to help (please tick one option)
you ask the application when you feel you need it ....
the application offers when it feels you need it........
both the above........
none of the above.......
14. Adaptation within a multimedia application (can take many forms) refers to the changing of links or content within the same application to different users i.e. users who are novice at using computers and the lecture content will have more help and more links to definitions etc whereas experts on computers and the content will be able to use all the system functions and will not have as many highlighted links to definitions (as it is assumed they already know them). If you were given the choice, which form of adaptation would you prefer (please tick as many as you like)
a. changing the presentation (i.e. the number of or the destination of) of highlighted links within a document...............
b. change the content of a document (by only displaying relevant parts of a document required to complete a task with irrelevant sections removed or shaded out)...............
c. allow access to all information available with all links possible (i.e. a conventional hypermedia teaching application).........

Why?..........

Appendix E First Empirical Study- Tasks for Content Novice User Group

Logging on information, task lists and application specific pre/ post session questionnaires for the Content Novice SM group.

Your status-Content Novice

Using the Adapted Navigation Application

Tasks

This application is based on Archaeology dating techniques. The links presented in this application are adapted to your knowledge of the subject domain.
Please read how to log onto the system and answer the questions.

1. Double click on the Microcosm Archaeology Dating Techniques icon
2. Log onto Microcosm using the user name 'ar_novice' and password 'novice', click on OK.
3. Then click on 'Dating techniques in Archaeology'and then on OK, now wait until a 'results box' and 'The Archaeology Dating techniques', 'Introduction Page' pop up.

5. Now read through the page and answer the questions relating to the various techniques in any order.

Notes
To navigate through the application double click on the highlighed text.
You do not need to use the menu

Dating Techniques	**Questions**
Obsidian	1. What are the number of sources of obsidian found in the Western Mediterranean
Radiocarbon	1. When using radio carbon dating what causes contamination in samples
Dendrochronology	1. How can information about environmental conditions be used to obtain information about the past
Archaeomagnetic	1. Name two requirements of archaeomagnetic dating

Post-Session Questionnaire

For the Adapted Navigation Application please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too many links			
there were too few links			
there were not enough links to definitions			
there were too many links to definitions			
did you need to revisit documents			
presentation of the documents was clear			
the documents contained too much text			
this was a useful form of adaptation			

Your status-Content Novice
Using the Adapted Content Application

Tasks

1. Log out of Microcosm
2. Log back into Microcosm by double clicking on the Microcosm Archaeology Dating Techniqües icon using the user name 'cont_ensn' and password 'contensn', click on OK.
3. Then double click on 'Archaeology (content)', now wait until a 'results box' and 'The Archaeology Dating techniques, Introduction Page' appears
4. Now read through the page and answer the questions relating to the various techniques in any order.
6. You are a novice on dating techniques in archealogy therefore please only answer the set questions using the text that has **not been shaded out**.

Now please complete the following task;

Dating Techniques	**Questions**
Obsidian	1. What are the magnetic properties of obsidian
Radiocarbon	1. How do you calculate a date using radio carbon dating
Dendrochronology	1. How is dendrochronology calculated
Archaeomagnetic	1. Out of radio carbon and magnetic dating which is more useful

Post-Session Questionnaire

For the Adapted Presentation Application please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too many links			
there were too few links			
there were not enough links to definitions			
there were too many links to definitions			
did you need to revisit documents			
presentation of the documents was clear			
the documents contained too much text			
this was a useful form of adaptation			
shading of text avoids information overload			
shaded text should be available to novice users			

Using the Control Application

Tasks

This application is the control application, it is presented as a conventional hypermedia application which allows the same access to all levels of users i.e.. links etc have not been adapted to your skill level

1. Log out of Microcosm
2. Log back into Microcosm by double clicking on the Microcosm Archaeology Dating Techniques icon using the user name 'control' and password 'control', click on OK.
3. Then double click on 'Archaeology (control)' and click on OK. now wait until a 'results box' and 'The Archaeology Dating techniques, Introduction Page' appears
4. Now read through the page and answer the questions relating to the various techniques in any order.

Now please complete the following task;

Dating Techniques	**Questions**
Obsidian	1. Can obsidian be used to date forgeries
Radiocarbon	1. Where did the 'Iceman' come from
Archeomagnetic	1. Can magnetic dating be used to date coins.

Post-Session Questionnaire

For the control application please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too many links			
there were too few links			
there were not enough links to definitions			
there were too many links to definitions			
did you need to revisit documents			
presentation of the documents was clear			
the documents contained too much text			
this was a useful form of adaptation			
do you think that adapting the content and the links in one application would be useful			
this application presented all links to all users was this useful in learning			

Appendix F First Empirical Study- Post-Session Questionnaire for the Content Novice User Group

Your status-Content Novice

Only answer this section after you have completed all the questions using both applications.

1. Adaptation within a multimedia application (can take many forms) refers to the changing of links or content within the same application to different users i.e. users who are novice at using computers and the lecture content will have more help and more links to definitions etc whereas experts on computers and the content will be able to use all the system functions and will not have as many highlighted links to definitions (as it is assumed they already know them).
   If you were given the choice, which form of adaptation would you prefer (please tick as many as you like)
   a. changing the presentation (1st application) (i.e. the number of or the destination of) of highlighted links within a document...............
   b. change the content of a document (2nd application) (by only displaying relevant parts of a document required to complete a task with irrelevant sections removed or shaded out)...............
   c. allow access to all information available with all links possible (3rd application) (i.e. a conventional hypermedia teaching application).........

Why?..
....

2. Please respond by ticking an answer under one of the headings;

	agree	not sure	disagree
computers are a good teaching tool			
adaptation is useful in teaching			
adaptation of links application required you to make less decisions compared to the control application			
adaptation of content application required you to make less decisions compared to the control application			

Appendix G First Empirical Study- Tasks for Content Expert User Group

Logging on information, task lists and application specific pre/ post session questionnaires for the Content expert user group.

Your status-Content Expert

Using the Adapting Links Application

Tasks

This application is based on Archaeology dating techniques. The links presented in this application are adapted to your knowledge of the subject domain.

Please read how to log onto the system and answer the questions.

1. Double click on the Microcosm Archaeology Dating Techniques icon
2. Log onto Microcosm using the user name 'ar_expert' and password 'expert', click on OK.
3. Then double click on 'Dating techniques in Archaeology', now wait until a 'results box' and 'The Archaeology Dating techniques, Introduction Page' appears

5. Now read through the page and answer the questions relating to the various techniques in any order.

Notes
To navigate through the application double click on the highlighted text.
You do not need to use the menu

Now please complete the following tasks in any order to help you the questions are written next to the relevant dating section;

Dating Techniques	**Questions**
Obsidian	1. How many sources of obsidian exist in the Western Mediterranean and name one special feature about 2 of the sources
	2. How is Fission Track Analysis used to date an artifact
Radiocarbon	1. How long does it take to measure a sample using tandom accelerators
Dendrochronology	1. Give 2 examples of how information about environmental conditions can be used to obtain information about the past
Archaeomagnetic	1. What is required to carry out archeomagnetic dating

Post-Session Questionnaire

For the Adapted Navigation Application- please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too many links			
there were too few links			
there were not enough links to definitions			
there were too many links to definitions			
did you need to revisit documents			

presentation of the documents was clear
the documents contained too much text
this was a useful form of adaptation

Your status-Content Expert

Using the Adapted Content Application

Tasks

Adaptation in this application is done by shading out parts of the text.

1. Log out of Microcosm
2. Log into Microcosm again by double clicking on the Microcosm Archaeology Dating Techniques icon using the user name 'contensn' and password 'contensn', click on OK.
3. Then double click on 'Archaeology (content)', now wait until a 'results box' and 'The Archaeology Dating techniques, Introduction Page' appears
4. Now read through the page and answer the questions relating to the various techniques in any order.

6. You are an expert on dating techniques in archaeology therefore please answer the set questions using **all the text including the text which has been shaded out.**

Now please complete the following task;

Dating Techniques	***Questions***
Obsidian	1. What does the value of magnetism present in obsidian depend on
Radiocarbon	1. What is the half life of radiocarbon in years
Dendrochronology	1. What are the problems can you face hwen calculating the date of timber
Archeaomagnetic	1. Out of radio carbon and magnetic dating which is more useful 2. At what point in time was it possible to gain high precision on magnetic dating

Post –Session Questionnaire

For the Adapted Presentation Application- please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too many links			
there were too few links			
there were not enough links to definitions			
there were too many links to definitions			
did you need to revisit documents			
presentation of the documents was clear			
the documents contained too much text			
this was a useful form of adaptation			
shading of text avoids information overload			
shaded text should be available to novice users			

Using the Control Application

Tasks

This application is the control; it is presented as a conventional hypermedia application, which allows the same access to all levels of users i.e.. links etc have not been adapted to your skill level .

1. Log out of Microcosm
2. Log onto Microcosm again by double clicking on the Microcosm Archaeology Dating Techniques icon using the user name 'control' and password 'control', click on OK.
3. Then double click on 'Archaeology (control)' and click on OK. now wait until a 'results box' and 'The Archaeology Dating techniques, Introduction Page' appears
4. Now read through the page and answer the questions relating to the various techniques in any order.

Now please complete the following task;

Dating Techniques	**Questions**
Obsidian	1. Can obsidian be used to detect forgeries
Radio Carbon	1. Which technique directly measures the ion count in radiocarbon 14
Dendrochronology	1. What effect would poor sunlight have on the growth of trees.
Archeomagnetic	1. What is the Curie point

Post-Session Questionnaire

For the control application- please answer the following questions after completing the above tasks by ticking under one heading;

	agree	not sure	disagree
you visited the right documents first time			
you easily found relevant documents			
there were too many links			
there were too few links			
there were not enough links to definitions			
there were too many links to definitions			
did you need to revisit documents			
presentation of the documents was clear			
the documents contained too much text			
this was a useful form of adaptation			
do you think that adapting the content and the links in one application would be useful			
this application presented all links to all users was this useful in learning			

Appendix H First Empirical Study- Post- Session Questionnaire for Content expert User Group

Post-Session Questionnaire

Your status-Content Expert

Answer this section after you have completed all the questions using both applications.

1. Adaptation within a multimedia application (can take many forms) refers to the changing of links or content within the same application to different users i.e. users who are novice at using computers and the lecture content will have more help and more links to definitions etc whereas experts on computers and the content will be able to use all the system functions and will not have as many highlighted links to definitions (as it is assumed they already know them).

   If you were given the choice, which form of adaptation would you prefer (please tick as many as you like)

   a. changing the presentation (i.e. the number of or the destination of) of highlighted links within a document..............

   b. change the content of a document (by only displaying relevant parts of a document required to complete a task with irrelevant sections removed or shaded out)...............

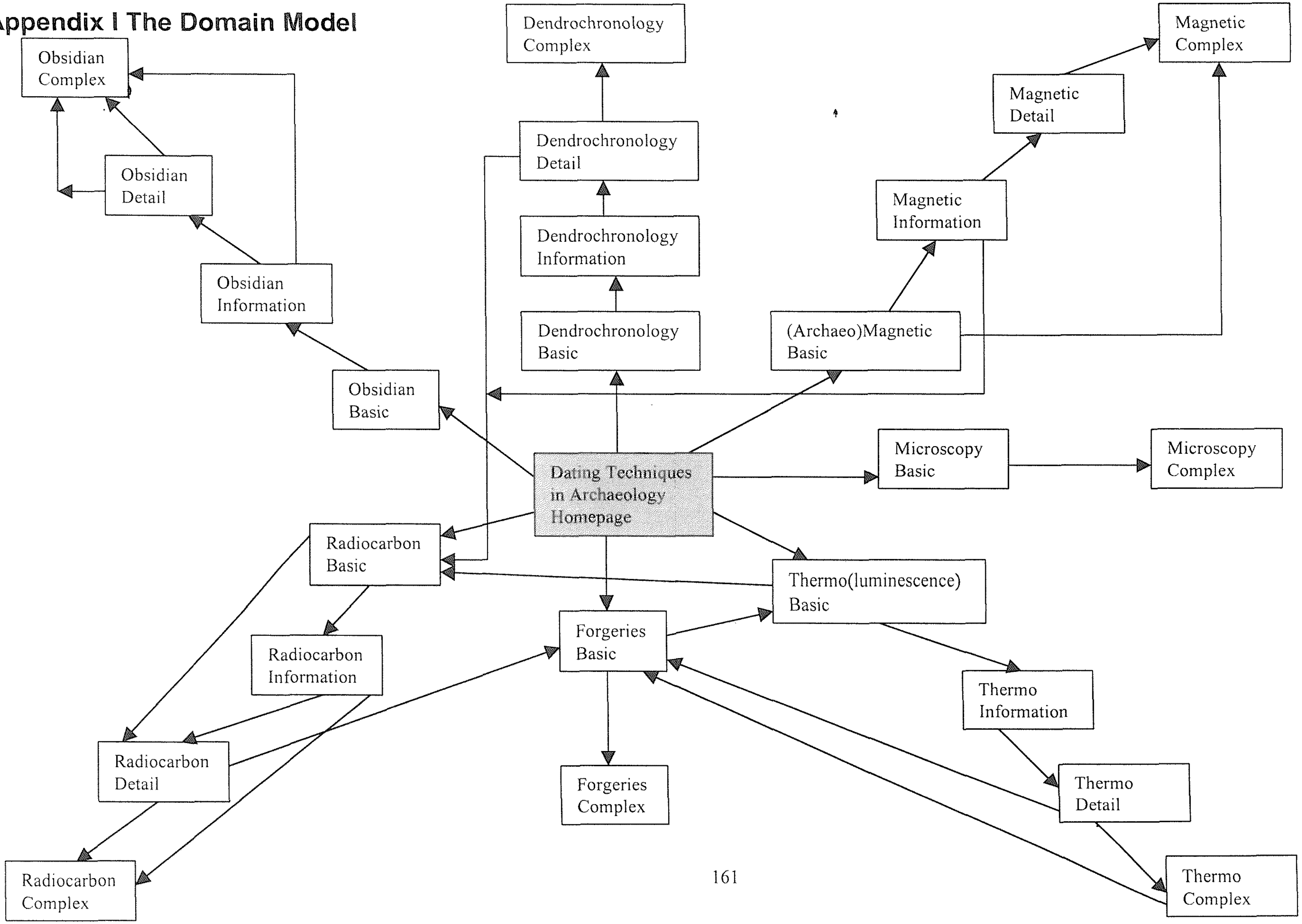
   c. allow access to all information available with all links possible (i.e. a conventional hypermedia teaching application).........

Why?..

2. Please respond by ticking an answer under one of the headings;

	agree	not sure	disagree
computers are a good teaching tool			
adaptation is useful in teaching			
adaptation of links application required you to make less decisions compared to the control application			
adaptation of content application required you to make less decisions compared to the control application			

Appendix I The Domain Model



Appendix J Second Empirical Study- Pre-Session Tasks

(These were also the tasks users were asked to answer during the evaluation sessions).
If you know the answer or think you know the answer please circle the relevant letter, A-D. If you do not know the answer, do not guess and circle the letter E.

Archaeomagnetic Dating

1. What materials are dated using Archaeomagnetic dating
   A) Clay walls
   B) Wood
   C) Stone
   D) Glass
   E) Do not know

2. What can upset the Archaeomagnetic qualities
   A) Rain
   B) Heat
   C) Carbon
   D) Snow
   E) Do not know

3. What are one of the benefits of Archaeomagnetic dating over radiocarbon dating
   A) Quicker to measure
   B) More accurate
   C) Less complex to measure
   D) Can measures a wider range of materials
   E) Do not know

4. What part of the kiln is it preferable to sample
   A) Walls
   B) Floor
   C) Ceiling
   D) Door
   E) Do not know

Dendrochronology

5. Age is calculated by
   A) Weighing the timber
   B) Counting the number of tree rings
   C) Calculating the carbon content
   D) Assessing environmental conditions over the possible time period
   E) Do not know

Microscopy

6. The polarizing microscope is used to identify
   A) Rocks
   B) Wood
   C) Metals
   D) Brick
   E) Do not know

Thermoluminescence

7. What materials are dated using thermoluminescence
   A) Obsidian
   B) Metal
   C) Wood
   D) Clay
   E) Do not know

8. Thermoluminescence is also useful in the detection of
   A) Metal
   B) Wood
   C) Ceramic artefacts
   D) Kiln walls
   E) Do not know

9. The principle behind Thermoluminescence is
   A) Measuring trapped energy
   B) Freezing the sample
   C) Measuring the magnesium content
   D) Measuring impurities of a sample
   E) Do not know

10. Energy is emitted in the form of
    A) Carbon
    B) Light
    C) Heat
    D) Radiocarbon
    E) Do not know

11. The energy is measured using a
    A) Mass Accelerator
    B) Photomultiplier
    C) Polarizing microscope
    D) Reference curve
    E) Do not know

12. The contamination problems which may occur in the soil are
    A) Radioactivity
    B) Too much exposure to light
    C) Water
    D) High temperatures
    E) Do not know

Radiocarbon Dating

13. What materials are measured by radiocarbon dating
    A) Metal
    B) Stone
    C) Bone
    D) Ceramics
    E) Do not know

14. A sample is contaminated if it
    A) Was contaminated after it was brought to the lab
    B) Contains at least 20% of roots/ stones
    C) Contains carbon which was not originally a component of the sample
    D) Contains deposits of metals and carbon which was an original component of the sample

E) Do not know

15. A tandom accelerator directly measures
    A) By weighing the components of the sample
    B) By ions counting
    C) The amount of light emitted
    D) The amount of contamination a sample contains
    E) Do not know

Obsidian Hydration Dating

16. How many groups have geologists divided obsidian into
    A) 4
    B) 5
    C) 3
    D) 2
    E) Do not know

17. How is the Hydration layer formed
    A) By contaminants
    B) By water
    C) Burning of an artefact
    D) By changes in environmental conditions
    E) Do not know

18. What equipment is used to measure the hydration layer
    A) Microscope
    B) Accelerator Mass Spectrometer
    C) Photomultiplier
    D) Filar micrometer
    E) Do not know

19. What can effect the hydration layer
    A) Soil
    B) Magnesium
    C) Wind
    D) Carbon
    E) Do not know

20. What percentage of water does the hydration layer contain
    A) 5%
    B) 3.5%
    C) 2.5%
    D) 10%
    E) Do not know

Appendix K Second Empirical Study- Pre-Session Questionnaire

Introduction to AH Adaptive hypermedia- to give user's some idea of what adaptive hypermedia is

Adaptive hypermedia refers to when a system adapts for example the content presented or the set of links presented to individual users based on the users task or knowledge levels. Examples of adaptive hypermedia include the following

Shading- where complex information is shaded out and can be ignored by novice users

produced accurate dates for larger timbers, thus helpir
archaeological sites.

Tree-ring calibration is important in the establishmen
and can sometimes be used in conjunction with other a
the interpretation of sites.

Now follow the next green link;

Changing the content presented depending on the knowledge of the user

Content expert users view

> However, the questions referring to the global maps showed marked differences between stages 2 and 3 . Whereas stage 3 users found it useful. Similarly in stage 2 50% disagreed with the statement' the complete map is unnecessary' compared

Content novice users view

> However, the questions referring to the global maps showed marked differences between stages 2 and 3 .

Changing the presentation of highlighted links to help the user to decide where to go next etc.

By changing the colour of the next relevant link

By hiding non relevant links

By annotating the highlighted links with different types of bullet points for example by indicating which link you are not yet ready for by an arrow

> ❖ Global maps
> ❖ Contents list
> ➢ Glossary

Pre Evaluation Questionnaire- Thank you for your co-operation These answers will remain confidential

	Question	**Answer** strongly agree 1	agree 2	not sure 3	disagree 4	strongly disagree 5
1	If you have taken an A'level in Archaeology please circle result	A B C D E F U Fail				
2	If you have taken an A'level in Archaeology please circle result	A B C D E F U Fail				
4	I have used educational computing applications before	1	2	3	4	5
5	I would not consider using a computer aided learning application	1				
6	I feel that a multiple number of windows within a learning package are useful	1	2	3	4	5
7	Adaptive hypermedia is a useful idea	1	2	3	4	5
8	I could easily define my content expertise levels on the online form	1	2	3	4	5
9	I would consider using a computer aided learning application to compliment lectures	1	2	3	4	5
10	*My current study aids are:*					
	Books	1	2	3	4	5
	Internet/ World Wide Web	1	2	3	4	5
	Computer aided learning applications	1	2	3	4	5
	Notes made by friends	1	2	3	4	5
	Notes produced myself from lectures	1	2	3	4	5
11	When using a CAL package I do not feel that I would need a tutor	1	2	3	4	5
12	When learning I prefer to use books because I can access any part of it even if I have not picked up enough knowledge about basic concepts	1	2	3	4	5
13	I prefer lectures because I am only exposed to information which I am currenlty ready to learn	1	2	3	4	5
14	When learning, information should not be made available if I am not yet ready to understand it	1	2	3	4	5
15	My level of computing experience is	1	2	3	4	5
16	I enjoy using computers when studying	1	2	3	4	5

Please place a tick in the relevant boxes

17	I use a computer for the following reasons and frequency	Never	Once a week	2-3 times a week	Daily
	E-mail				
	Internet/ WWW				
	Word processor				
	Computer games				
	Graphics package				

	My current knowledge on the following topics is	No knowledge	Basic Knowledge	More detail	Great deal of detail
	Radiocarbon dating Basic definition of radiocarbon dating How to collect samples AMS				
	Obsidian (Fissian track analysis) FTA What obsidian is FTA				
	Dendrochronology What Dendrochronology is				
	Thermoluminescence What Thermoluminescence is				
	Potassium Argon What Potassium Argon is				

Appendix L Second Empirical Study- Post-Session Questionnaire for the Users of Application A

Post Evaluation Usability Questionnaire For Application A users. Thank you for your co-operation
These answers will remain confidential

	Question	Answer strongly agree 1	agree 2	not sure 3	disagree 4	strongly disagree 5
Subjective satisfaction- your personal attitude towards the application						
1	I would consider using this application when studying	1	2	3	4	5
2	I feel that a multiple number of windows within a learning package are useful	1	2	3	4	5
3	Adaptive hypermedia is a useful idea	1	2	3	4	5
4	I could easily define my content expertise levels on the online form	1	2	3	4	5
5	I would use an application such as this in place of a teacher	1	2	3	4	5
6	It was easy to distinguish links I had already visited					
7	It was easy to distinguish the next relevant links					
8	It was easy to distinguish reference links i.e. links to glossary terms and images.					
9	I found relevant information quickly and easily					
10	I prefer lecture type teaching/ learning sessions					
11	I found the application confusing to use					
12	I talked about the application/ questions with my friends					
13	I did not need to ask the help of my tutor during the lab session					
14	This application was pleasant to use	1	2	3	4	5
Learnability-your ability to pick up relevant information efficiently						
15	I found it necessary to revisit documents	1	2	3	4	5
16	I feel that I learnt a great deal about archaeological dating techniques by using the application	1	2	3	4	5
17	I feel that I learnt the basics of archaeological dating techniques by using the application	1	2	3	4	5
18	I prefer this type of learning application to paper based learning techniques	1	2	3	4	5
19	I could easily get to the information I required	1	2	3	4	5
Ease of Use- your ability to easily move around the system and learn the basic structure of the hypertext network.						
20	The table of contents at the end of each document provided a good method of moving around the application	1	2	3	4	5
21	The table of contents provided me with some idea of the structure of the hypertext document network	1	2	3	4	5
22	I found that using the table of contents to move around the application was helpful	1	2	3	4	5
23	I found that having links within the text sections of the document and the table of contents was very helpful	1	2	3	4	5
24	I found the links within the text section of the document confusing	1	2	3	4	5
25	The dynamic change of link colours lead to confusion	1	2	3	4	5
26	Shading parts of the text disrupted my reading of the text	1	2	3	4	5
27	I knew which links to follow next	1	2	3	4	5
28	The guidance given by the application was useful	1	2	3	4	5
29	This application was generally confusing to use	1	2	3	4	5
30	This application maintained too many adaptive hypermedia techniques	1	2	3	4	5
31	The guidance given by the coloured links was very useful	1	2	3	4	5

32	I often became lost and disoriented when using the application	1	2	3	4	5
Flexibility- you are able to customise the adaptation of the system to your requirements						
33	I felt I had control over the system	1	2	3	4	5
34	Understanding the adaptability of the application was easy.	1	2	3	4	5
35	The guidance given by shading sections of text was very useful	1	2	3	4	5
36	I followed the green highlighted links because they were helpful	1	2	3	4	5
37	I sometimes followed the red links knowing I was not supposed to	1	2	3	4	5
38	The documents presented were at the right level for my knowledge	1	2	3	4	5

	My current knowledge on the following topics is	No knowledge	Basic Knowledge	More detail	Great deal of detail
	Radiocarbon dating Basic definition of radiocarbon dating How to collect samples AMS				
	Obsidian (Fissian track analysis) FTA What obsidian is FTA				
	Dendrochronology What Dendrochronology is				
	Thermoluminescence What Thermoluminescence is				
	Potassium Argon What Potassium Argon is				

Appendix M Second Empirical Study- Post-Session Questionnaire for the Users of Application B

Post Evaluation Usability Questionnaire For the Application B Users. - Thank you for your co-operation
These answers will remain confidential

	Question	Answer: strongly agree 1	agree 2	not sure 3	disagree 4	strongly disagree 5
Subjective satisfaction- your personal attitude towards the application						
1	I would consider using this application when learning	1	2	3	4	5
2	I feel that a multiple number of windows within a learning package are useful	1	2	3	4	5
3	Adaptive hypermedia is a useful idea	1	2	3	4	5
4	I could easily define my content expertise levels on the online form	1	2	3	4	5
5	I would use an application such as this in place of a teacher	1	2	3	4	5
6	This application was pleasant to use	1	2	3	4	5
Learnability-your ability to pick up relevant information efficiently						
7	I found it necessary to revisit documents	1	2	3	4	5
8	I feel that I learnt a great deal about archaeological dating techniques by using the application	1	2	3	4	5
9	I feel that I learnt the basics of archaeological dating techniques by using the application	1	2	3	4	5
10	I prefer this type of learning application to paper based learning techniques	1	2	3	4	5
11	I could easily get to the information I required	1	2	3	4	5
Ease of Use- your ability to easily move around the system and learn the basic structure of the hypertext network						
12	The table of contents at the end of each document provided a good method of moving around the application	1	2	3	4	5
13	The table of contents provided me with some idea of the structure of the hypertext document network	1	2	3	4	5
14	I found that using the table of contents to move around the application was helpful	1	2	3	4	5
15	I found that having links within the text sections of the document and the table of contents was very helpful	1	2	3	4	5
16	I found the links within the text section of the document confusing	1	2	3	4	5
17	The dynamic change of link colours lead to confusion	1	2	3	4	5
18	I knew which links to follow next	1	2	3	4	5
19	The guidance given by the application was useful	1	2	3	4	5
20	This application was generally confusing to use	1	2	3	4	5
21	This application maintained too many adaptive hypermedia techniques	1	2	3	4	5
22	The guidance given by the coloured links was very useful	1	2	3	4	5
23	I often became lost and disoriented when using the application	1	2	3	4	5
Flexibility- you are able to customise the adaptation of the system to your requirements						
24	I felt I had control over the system	1	2	3	4	5
25	Understanding the adaptability of the application was easy.	1	2	3	4	5
26	I followed the green highlighted links because they were helpful	1	2	3	4	5
27	I sometimes followed the red links knowing I was not supposed to	1	2	3	4	5
28	The documents presented were at the right level for my knowledge	1	2	3	4	5

	My current knowledge on the following topics is	No knowledge	Basic Knowledge	More detail	Great deal of detail
	Radiocarbon dating Basic definition of radiocarbon dating How to collect samples AMS				
	Obsidian (Fissian track analysis) FTA What obsidian is FTA				
	Dendrochronology What Dendrochronology is				
	Thermoluminescence What Thermoluminescence is				
	Potassium Argon What Potassium Argon is				

Appendix N Second Empirical Study- Post-Session Questionnaire for the Users of Application C

Post Evaluation Usability Questionnaire For Application C. - Thank you for your co-operation These answers will remain confidential

	Question	Answer				
		strongly agree 1	agree 2	not sure 3	disagree 4	strongly disagree 5
Subjective satisfaction- your personal attitude towards the application						
1	I would consider using this application when learning	1	2	3	4	5
2	I feel that a multiple number of windows within a learning package are useful	1	2	3	4	5
3	I could easily define my content expertise levels on the online form	1	2	3	4	5
4	I would use an application such as this in place of a teacher	1	2	3	4	5
5	This application was pleasant to use	1	2	3	4	5
Learnability-your ability to pick up relevant information efficiently						
6	I found it necessary to revisit documents	1	2	3	4	5
7	I feel that I learnt a great deal about archaeological dating techniques by using the application	1	2	3	4	5
8	I feel that I learnt the basics of archaeological dating techniques by using the application	1	2	3	4	5
9	I prefer this type of learning application to paper based learning techniques	1	2	3	4	5
10	I could easily get to the information I required	1	2	3	4	5
Ease of Use- your ability to easily move around the system and learn the basic structure of the hypertext network						
11	The table of contents at the end of each document provided a good method of moving around the application	1	2	3	4	5
12	The table of contents provided me with some idea of the structure of the hypertext document network	1	2	3	4	5
13	I found that using the table of contents to move around the application was helpful	1	2	3	4	5
14	I found that having links within the text sections of the document and the table of contents was very helpful	1	2	3	4	5
15	I found the links within the text section of the document confusing	1	2	3	4	5
16	I knew which links to follow next	1	2	3	4	5
17	The guidance given by the application was useful	1	2	3	4	5
18	This application was generally confusing to use	1	2	3	4	5
19	I often became lost and disoriented when using the application	1	2	3	4	5
Flexibility- you are able to customise the adaptation of the system to your requirements						
20	I felt I had control over the system	1	2	3	4	5
21	The documents presented were at the right level for my knowledge	1	2	3	4	5

	My current knowledge on the following topics is	No knowledge	Basic Knowledge	More detail	Great deal of detail
	Radiocarbon dating Basic definition of radiocarbon dating How to collect samples AMS				
	Obsidian (Fissian track analysis) FTA What obsidian is FTA				
	Dendrochronology What Dendrochronology is				
	Thermoluminescence What Thermoluminescence is				
	Potassium Argon What Potassium Argon is				

Appendix O Survey Data

237 subjects took part in the first survey. The following tables show the results. Please note that some subjects did not answer all of the questions.

4.4.4.1

Confidence on computers by Gender

	Male %	Female %
Computer Confident	52	21
Neutral	29	65
Not Computer Confident	7	7

Confidence at using e-mail by Gender

	Male %	Female %
e-mail expert	20	11
Neutral	58	46
e-mail novice	16	36

Confidence at using MS Office by Gender

	Male %	Female %
MS Office expert	17	8
Neutral	58	74
MS Office novice	19	14

Confidence at using the WWW by Gender

	Male %	Female %
WWWW expert	21	4
Neutral	41	41
WWW novice	33	50

MCM Windows Freedom by Computer Confidence

	Confident at using a computer (numbers)	Neutral (numbers)	Not confident at using a computer (numbers)
MCM allows too much windows freedom	29	43	2
Neutral	33	37	5
MCM does not allow too much windows freedom	2	5	2

Prefer the system to provide hints as assistance by windows management experience

	Expert at windows management (%)	Neutral (%)	Expert at windows management (%)
Require hints from the system	23	33	43

Windows management experience by Gender

	Expert at windows management %	Neutral %	Expert at windows management %
Male	21	38	39
Female	13	41	44

4.4.4.2

	Computer confident %			e-mail %			Word processor %			WWW %		
	Exp	Neu	Nov	Exp	Neu	Nov	Exp	Neu	Nov	Exp	Neu	Nov
Computer science	45	55	0	55	45	0	38	55	5	38	62	0
Arts	13	60	26	13	73	13	26	60	13	30	40	30
Medicine				11	45	44	38	50	12	13	21	63
Oceanography	39	36	21	17	78	5	19	75	5	4	75	17
Mechanical engineering	50	25	7	21	67	11	21	67	11	23	50	28

Arts include subjects from Archaeology, Language Centre and History.

4.4.4.3

Not all of the subjects answered this question

User Prefers to ask the application for assistance by found relevant information easily

	Found relevant information easily (%)	Neutral (%)	Found relevant information easily (5)
Prefer to ask application for assistance	42	39	19

User Prefers to ask the application for assistance by found relevant information easily

	Computer confident (%)	Neutral (%)	Not computer confident (numbers)
Prefer application to make changes itself automatically	39	36	-
Do not prefer application to make changes itself automatically	61	64	-

4.4.4.4

User gets distracted by age group

Age Group	Easily Distracted %	Neutral %	Not easily Distracted %
1-2	82	74	0
3-4	18	23	0
5-6	0	3	0

4.4.4.5

Country of High School Education	Computer Confident %	Neutral %	Not Computer Confident %
Europe	84	91	29
America	0	1	1
Far East in Malaysia, China, Hong Kong, Singapore	7	12	0
Africa and India inc. Sudan, Nigeria	4	0	1

References

Akoulchina, I and Ganascia, J., 1997. SATELIT- Agent: An Adaptive Interface Based on Learning Interface Agents Technology in User Modelling. *In: Conference Proceedings of the 6th International Conference UM 1997.* pp 21-32, Springer Wien.

Bailey, C. and Hall, W., 2000. An Agent Based Approach to Adaptive Hypermedia Using a Link Service. *In: Conference Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web-based Systems, AH 2000,* Trento, Italy, pp 260-263.

Beaulieu, M. and Mellor, V., 1995. The Micro Gallery: An Evaluation of the Hypertext System in the National Gallery, London. *In: The New Review of Hypermedia and Multimedia.* pp 233-260.

Beaumont, I., 1994. User Modelling and Hypertext Adaptation. *In: User Modelling and User Adapted Interaction.* Vol. 4 No.1, pp21-45.

Beaumont, I. and Brusilovsky, P., 1995. Adaptive Educational Hypermedia: From Ideas to Real Systems. *In*: Maurer, H, ed. *Conference Proceedings Ed-Media 1995.* AACE, pp 93-98.

Beck, J. Stern, M. and Woolf, B., 1997. Using the Student Model to Control Problem Difficulty in User Modelling. *In: Conference Proceedings of the 6th International Conference UM 1997.* pp 277-288, Springer Wien.

Benaki, E, Karkaletsis, V. and Spyropoulos, C., 1997. Integrating User Modelling Into Information Extraction: The UMIE Prototype, User Modelling. *In: Conference Proceedings of the 6th International Conference, UM 1997.* pp 55-57, Springer Wien.

Benyon, D. and Murray, D., 1988. Experience with Adaptive Interfaces. *In: The Computer Journal*, Vol. 31 No. 5. pp465-473.

Boyle, C. and Encarnacion, A. O., 1994. MetaDoc: An Adaptive Hypertext Reading System. *In: User Modelling and User Adapted Interaction*, Vol. 4 No. 1, pp1-19.

Brusilovsky, P., 1992. Intelligent Tutor, Environment and Manual for Introductory Programming. *In: Educational and Training Technology International,* Vol. 29 No. 1, pp26-34.

Brusilovsky, P., 1994 (b) Adaptive Hypermedia: An Attempt to Analyse and Generalise, *In: Adaptive Hypermedia Workshop, UM 1994, MA, USA.*

Brusilovsky, P. and Eklund, J., 1998 (a). A Study of User Model Based Link Annotation in Educational Hypermedia. I*n: Journal of Universal Science,* Vol. 4 No. 4. pp429-448.

Brusilovsky, P., Kobsa, A., and Vassileva, J., 1998 (b). *Adaptive Hypertext and Hypermedia,* Dordrecht: Kluwer Academic Publishers.

Brusilovsky, P. and Cooper, D., 1999. ADAPTS: Adaptive Hypermedia for a Web-based Performance Support System. *In:* P. Brusilovsky and P. De Bra eds. *Conference Proceedings of the 2nd Workshop on Adaptive Systems and User Modelling on the WWW* at the 8th International WWW Conference.

Brusilovsky, P. and Pesin, L., 1994. ISIS-TUTOR: An Intelligent Learning Environment for CDS/ISIS Users. *In:* J. Levonon and M. Tukianinen eds., *Proceedings of the Interdisciplinary Workshop on Complex Learning in Computer Environments,* pp 29-33.

Brusilovsky, P. and Pesin, L., 1995 (b). Visual Annotation of Links in Adaptive Hypermedia *In:* I. Katz, R. Mack and L. Marks eds. *Conference Proceedings CHI 1995,* pp 222-223, ACM Publishers.

Bull, S., 1997. See Yourself Write: A Simple Student Model to Make Students Think. *In: Conference Proceedings of the 6th International Conference on UM 1997.* pp 339-341, Springer Wien.

Bush, V., 1945. As We May Think. *In: Atlantic Monthly.* http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm.

Cakmakov, D. and Davcev, D., 1993. An Intelligent User Interface for Multimedia Mineral Retrieval System in Human Computer Interaction: Software and Hardware Interfaces. *In:* G. Salvendy and M. Smith, eds. *Conference Proceedings of the 5th International Conference on HCI*, Vol. 2. Elsevier.

Carr, L., DeRoure, D., Davis, H., Hall, W., 1998. Implementing an Open Link Service for the World Wide Web. *In: World Wide Web Journal 1998,* Vol. No. 2, pp 61-71.

Chanier, T., 1996. Evaluation as Part of the Project Lifecycle: The Hypermedia Camille Project. *In: Journal of the Association for Learning Technology*, Vol. 4 No. 3. pp 54-68

Cheverst, C., Davies, N., Mitchell, K. and Smith, P., 2000. Providing Tailored (Context-Aware) Information to City Visitors. *In:* P. Brusilovsky, O. Stock and C. Strapparava, eds. *Lecture Notes in Computer Science, Adaptive Hypermedia and Adaptive Web-Based Systems, Conference Proceedings AH200*, Trento, Italy, pp 73-85, Springer.

Cockton, G., 1986. Some Critical Remarks on Abstractions for Adaptable Dialogue Managers. *In:* M. Harrison and A. Monk, eds. *BCS Workshop Series, People and*

Computers: Designing for Usability, Conference Proceedings of 2nd Conference of BCS HCI Specialist Group. Cambridge University Press.

Cohen, A. and Calvacanti, M., 1990. Feedback on Composition. *In: Teacher and Student Verbal Reports in Kroll.*

Cooper, M., 1988. Interfaces that Adapt to the User. *In: Artificial Intelligence and Human Learning: Intelligent Computer Aided Instruction.* Chapman and Hall.

Da Silva, D. P., 1998. Concepts and Documents for Adaptive Educational Hypermedia: A Model and a Prototype. *In: Computing Science Reports.* Eindhoven University of Technology, pp 33-41.

Davis, H., Hall, W., Heath, I. and Hill, G., 1992. Towards an Integrated Information Environment with Open Hypermedia Systems. *In:* D. Lucarella, J. Namard, M. Namard and P. Paolini, eds. *Conference Proceedings of the ACM conference on Hypertext.* ACM Press.

De Bra, P. and Calvi, L., 1998 (a). AHA: A Generic Adaptive Hypermedia System. *In: Proceedings of the 2nd Workshop on Adaptive Hypertext and Hypermedia, Hypertext Conference 1998*, Pittsburgh, pp 5-11.

De Bra, P. and Calvi, L., 1998 (b). AHA! An Open Adaptive Hypermedia Architecture *In: The New Review of Hypermedia and Multimedia*, Vol. 4, pp 115-139, Taylor Graham Publishers.

De Bra, P., Houben, G. and Wu, H., 1999, AHAM: A Dexter Based Reference Model for Adaptive Hypermedia. *In: Conference Proceeding of Hypertext 1999*, Darmstadt, pp 147-156, ACM.

Desurvire, H. W., 1994. Faster, Cheaper! Are Usability Inspection Methods as Effective as Empirical Testing? *In:* J. Nielsen and Mack, R. L., eds. *Usability Inspection Methods,* John Wiley & Sons, Inc.

Diaper, D., 1987. Identifying the Knowledge Requirements of an Expert System's Natural Language Processing Interface. *In:* D. Diaper and R. Winder, eds. *Conference Proceedings of 3rd Conference of BCS HCI Specialist People and Computers.* pp 263-286, Cambridge University Press.

Dix, A., Finlay, J., Abowd, G. and Beale, R., 1993. *Human Computer Interaction,* Prentice Hall International (UK) Ltd.

Dufresne, A., Adaptive Hypermedia: Supporting the Communication Process.

Eklund, J. and Brusilovsky, P., 1998 (a). The Value of Adaptivity in Hypermedia Learning Environments: A Short Review of Empirical Evidence. *In: Computing Science Reports.* No. 98/12, Eindhoven University of Technology, Eindhoven, pp 13-19.

Eklund, J. and Brusilovsky, P., 1998 (b). Individualising Interaction. In: Web-based Instructional Systems in Higher Education. *In: Conference Proceedings of the AUC Academic Conference*, Melbourne, Austrailia, pp 27-30.

Eklund, J., Brusilovsky, P., and Schwarz, E., 1998 (c). A Study of Adaptive Link Annotation. *In* T. Ottmann and I. Tomek eds. *Conference Proceedings of EDMEDIA/ ED-TELECOM 1998.* pp 304-309, Springer Online, Germany.

Elsom-Cook, M., 1988, Guided Discovery Tutoring and Bounded User Modelling. *In* J. Self ed. *AI and Human Learning.* pp- 165-178, Chapman and Hall.

Encarnacao, L., 1997, Multi Level User Support Through Adaptive Hypermedia: A Highly Application Independent Help Component. *In:* eds. J. Moore, E. Edmonds and A. Puerta, *Conference Proceedings Intelligent User Interfaces 1997*, ACM, pp 187-194.

Espinoza, F. and Hook, K., 1996. A WWW Interface to an Adaptive Information System *In: Workshop User Modelling for Information Filtering on the WWW Conference Proceedings of the 5th International Conference on UM 1996.*

Fink, J. and Kobsa, A., 1998. Adaptable and Adaptive Information Provision for all Users, Including Disabled and Elderly People. *In: The New Review of Hypermedia and Multimedia* Vol. 4. pp 163-188.

Fowler, C., Macaulay, L. and Siripoksup, S., 1986. An Evaluation of the Effectiveness of the Adaptive Interface Module (AIM). *In:* M. Harrison and A. Monk, eds. *Matching Dialogues to Users in BCS Workshop Series People and Computers Designing for Usability Conference Proceedings of the 2nd BCS HCI Specialist Group.* pp 345-359, Cambridge University Press.

Gillham, M., Kemp, B. and Buckner, K., 1995. Evaluating Interactive Multimedia Product for the Home. *In: New Review of Hypermedia*, Vol. 1. pp199-212.

Gutkauf, B., Thies, S. and Domik, G., 1997. A User-Adaptive Chart Editing System Based on User Modelling and Critiquing in User Modelling. *In: Conference Proceedings of the 6th International Conference UM 1997.* pp159-170, Springer Wien.

Hall, W., Weal, M., Heath, I., Wills, G. and Crowder, R., 1997. Flexible Interfaces in the Industrial Environment. *In: International Conference Managing Enterprises-Stakeholders, Engineering, Logistics and Achievement (ME-SELA '97).*

Hall, W., 2000. As We May Learn- The Application of Hypermedia in Education. *In: Computing and control Engineering Journal,* pp173-178.

Hill, L., Dolin, R., Frew, J. Kemp, R. Larsgaard, M., Montello, D., Rae, M. and Simpson, J., 1997. User Evaluation: Summary of the Methodologies and Results for the Alexandria Digital Library, University of California at Santa Barbara. http://www.asis.org/annual-97/alexia.htm

Hix, D. and Harrison, H. R., 1993. *Developing User Interfaces.* John Wiley & Sons.

Hohl, H., Bocker, H. and Gunzenhauser, R., 1996. Hyperadapter: An Adaptive Hypertext System for Exploratory Learning and Programming *In: User Modelling and User Adapted Interaction,* No. 6. pp 131-156, Kluwer Academic Publishers.

Hook, K., 1997. Evaluating the Utility and Usability of an Adaptive Hypermedia System. *In:* J. Moore, E. Edmonds and A. Puerta, eds. *Conference Proceedings Intelligent User Interfaces.* ACM. pp 179-186.

Hutchings, G., Hall, W., Briggs, J., Hammond, N. V. Kibby, M. R., McKnight, C. and Riley, D., 1992. Authoring and Evaluation of Hypermedia for Education. *In: Computers Education,* Vol. 18 No. 1-3. pp 171-177, Pergamon Press plc.

Hutchings, G., Hall, W. and Thorogood, P., 1994. Experiences with Hypermedia in Undergraduate Education. *In: Computers in Education,* Vol. 22, No. 1. Pergamon Press.

Jacobs, G., 1998. Evaluating Courseware: Some Critical Questions. *In: Innovations in Education and Training International,* Vol 35, No.1. pp 3-8.

Johnson, P., 1992. Human Computer Interaction, McGraw Hill,[pp235-239].

Jordan, P., Draper, S., MacFarlane, K. and McNulty, S., 1991. Guessability, Learnability and Experienced User Performance. *In: D.* Diaper and N. Hammond, Eds. *BCS Conference Series People and Computers V1 Conference Proceedings of the HCI 1991 Conference.* Cambridge University Press.

Jorgensen, A. H. 1990., Thinking Aloud in User Interface Design: A Method Promoting Cognitive Ergonomics. *In: Ergonomics*, Vol 33, No. 4, pp 501-507.

Kalyuga, S., Chandler, P and Sweller, J, 1997. Levels of expertise and User Adapted Formats of Instructional Presentations: A Cognitive Approach. *In: Conference Proceedings of the 6thy International Conference on User Modelling.* pp 261-272, Springer Wien.

Kaplan, C., Fenwick, J. and Chen, J., 1993 Adaptive Hypertext Navigation Based on User Goals and Context. *In: Adaptive Hypertext and Hypermedia,* 1998, Dordrecht: Kluwer Academic Publishers.

Karat, C., Campbell, R. and Fiegel, T., 1992. Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation in Human Factors. *In:* P. Bauersfeld, J. Benneett and G. Lynch, eds. *Conference Proceedings Computing Systems - Striking a Balance.* pp 397-404, ACM Press, Addison Wesley.

Kato, T., 1986. What Question Asking Protocols can Say About the User Interface. *In: International Journal of Man-Machine Studies*, Vol. 25, pp659-673.

Kemp, M., 2000. From Classroom Tutor to Hypermedia Advisor; A Case Study in Medical Education, PhD Thesis, University of Southampton, June 2000.

Kornbrot, 1990. Monitoring and Analysis of Hypermedia Navigation. *In:* D. Diaper, ed. *HCI Interact,* 1990. pp 401-406.

Kobsa, A., Muller, D. and Nill, A., 1994. KN-AHS: An Adapyive Hypertext Client of the User Modelling System BGP-MS. *In: Conference Proceedings. Of the 4th International Conference on User Modelling*, pp31-36.

Laurrilard, D., 1998. Multimedia and the Learners Experience of Narrative. *In Computers in Education,* Vol. 31, No.2.

Lea, M., 1998. Evaluating User Interface Designs. *In:* T. Rubin ed. *User Interface Design for Computer Systems,* pp134-167, Ellis Horwood.

Linton, F., Joy, D. and Schaefer, H., 1999. Building User and Expert Models by Long-Term Observation of Application Usage. *In* J. Kay ed. *Conference Proceedings. Of 7th International Conference on User Modelling.* pp 129-138, Springer, New York.

Makrakis, V., Retalis, S., Koutoumanos, A., Papaspyrou, N. and Skordalakis, M., 1998. Evaluating Effectiveness of an ODL Hypermedia System and Courseware at the National Technical University of Athens: A Case Study *In: JUCS* Vol. 4, Issue 3. pp259-272.

Malinowski, U., Kuhme, T., Dieterich, H. and Schneider-Huschmidt, M., 1993. Computer-Aided Adaptation of User Interfaces With Menu's and Dialogue Boxes. *In:* G. Salvendy and M. Smith, eds. *Human Computer Interaction: Software and Hardware Interfaces Conference Proceedings of 5th International Conference on HCI,* Vol. 2. Elsevier.

Marchionini, G. and Crane, H., 1994. Evaluating Hypermedia and Learning Methods and Results form the Perseus Project. *In: ACM Transactions on Information Systems,* Vol. 12, No. 1. pp5-34.

Maulsby, D., 1997. Inductive Task Modelling for User Interface Customisation. *In:* J. Moore, E. Edmonds and A. Puerta, eds. *Intelligent User Interface Conference Proceedings 1997.* pp 233-236, ACM Press.

Molich, R. and Nielsen, J., 1990. Improving a Human-Computer Dialogue. In : Communications of the ACM, Vol. 33, No. 3. pp 338-348.

Morris, A., 1986. Expert Systems- Interface Insight. *In:* M. Harrison and A. Monk, eds. *Matching Dialogues to Users in BCS Workshop Series People and Computers Designing for Usability Conference Proceedings of 2nd Conference of BCS HCI Specialist Group.* Cambridge University Press.

Murphy, M. and McTear, M., 1997. Learner Modelling for Intelligent CALL. *In:* A. Jameson, C. Paris and C. Tasso, eds. *User Modelling; Conference Proceedings of the 6th International Conference.* pp 301-312, Springer Wien.

Nelson, T., 1980. Replacing the Printed Word: A Complete Literary System. *In:* S. H. Lavington ed. *Proceedings of the IFIP Congress* pp 1013-1023.

Nemetz, F., Winckler, M. A. and de Lima, J. V., 1997. Evaluating Evaluation Methods For Hypermedia Applications. *In:Conference Proceedings. Ed-Media 1997.*

Newman, W. and Lamning, M., 1995. *Interactive System Design.* Addison Wesley.

Nicol, A., 1990. Interfaces For Learning: What Do Good Teachers Know That We Don't. *In:* B. Laurel, ed. *The Art of Human Computer Interface Design.* Addison Wesley.

Nielsen, J., 1990. Evaluating Hypertext Usability. *In:* D. H. Jonassen and H. Mandl, eds. *Designing Hypermedia for Learning.* NATO ASI Series. Vol 67, pp 147-168, Springer-Verlag, Berlin.

Nielsen, J., 1992. Finding Usability Problems Through Heuristic Evaluation. *In:* P Bauersfeld, J. Bennett and G. Lynch, eds. *ACM Conference on Human Factors in*

Computing Systems, Conference Proceedings CHI 1992. pp 373-380, ACM Press, Addison Wesley.

Nielsen, J., 1993. *Usability Engineering*, Academic Press.

Nielsen, J. and Morlich, R., 1990. Heuristic Evaluation of User Interfaces. *In:* J. Carrasco, J. Whiteside, eds. *Human Factors in Computing Systems Empowering People* . pp 249-256, ACM Press.

Nill, A., 1997. Providing Useable and Useful Information by Adaptivity and Adaptability. *In: Proceedings of the Flexible Hypertext Workshop, 8th International Hypertext Conference,* Southampton, UK.

Orsini-Jones, M. and Jones, D., 1996. Hypermedia for Language Learning: the FREE Model at Coventry University. *In: Association for Learning Technology Journal,* Vol. 4, No. 2. pp 28-39.

Orwant, J., 1991. The Doppelganger User Modelling System. *In: Proceedings of the IJCAI Workshop W4, Agent Modelling for Intelligent Interaction,* Sydney, Australia, pp 164-168.

Perez-Quinones, M. and Sibert, J., 1996. A Collaborative Model Of Feedback in Human-Computer Interaction. *In: Conference Proceedings CHI 1996,* ACM. pp 316-323.

Plessis, J., Van Biljon, A., Tolmie, c and Wollinger, T, 1995. A Model For Intelligent Computer-Aided Education Systems. *In: Computers in Education,* Vol. 24, No. 2. pp 363-368, Elsevier Siences Ltd.

Rafter, R., Bradley, K. and Smyth, B., 2000. Automated Collaborative Filtering Applications for Online Recruitment Services. *In:* P. Brusilovsky, O. Stock and C.

Strapparava, eds. *Lecture Notes in Computer Science, Adaptive Hypermedia and Adaptive Web-Based Systems, Conference Proceedings AH200*, Trento, Italy, Springer.

Recker, M., 1994. A Methodology for Analysing Students' Interactions within Educational Hypermedia. *In: Educational Multimedia and Hypermedia Annual.*

Rouet, J. F., 1992. Cognitive Processing of Hyperdocuments: When Does Nonlinearity Help?. *In: Conference Proceedings ACM ECHT*, Milano, Nov 30- Dec 4.

Salzman, M. C. and Rivers, S. D., 1994. Smoke and Mirrors: Setting the Stage for a Successful Usability Test. *In: Behaviour & Information Technology*, Vol 13, No. 1 and 2, pp9-16.

Sandberg, J., 1987. *In: Conference Proceedings of the 3rd International Conference on Artificial Intelligence and Education, AICOM.*

Schlingbaum, E., 1997. Individual User Interfaces and Model-based User Interface Software Tools. *In:* J. Moore, E. Edmonds and A. Puerta, eds. *Intelligent User Interfaces Conference Proceedings 1997.* pp 229-232, ACM Press.

Seta, K., Ikeda, M., Kakusho, O. and Mizoguchi, R., 1997. Capturing a Conceptual Model for End-User Programming: Task Ontology as a Static User Model. *In: Conference Proceedings. 6th International Conference on UM 1997,* pp 203-214.

Shneiderman, B., 1997. Direct Manipulation for Comprehensible, Predictable and Controllable User Interfaces. *In:* J. Moore, E. Edmonds and A. Puerta, eds. *International Conference Proceedings on Intelligent User Interfaces.* ACM, pp 33-39.

Spall, R. and Steele, R., 1990. An Investigation Into Quantitative User Modelling of User Interactions For The Purpose of Predicting User Expertise. *In: D. Diaper, ed. Human-Computer Interaction-INTERACT'90.* pp 129-134, Elsevier Science.

Strachan, L., Anderson, J., Sneesby, M. and Evans, M., 1997. Pragmatic User Modelling in a Commercial Software System. *In:* A. Jameson, C. Paris and C. Tasso, eds. *Conference Proceedings of the 6th International Conference on UM 1997.* pp 189-200, Springer Wien.

Stern, M. and Park-Woolf, B., 2000. Adaptive Content in an Online Lecture System. *In:* P. Brusilovsky, O. Stock and C. Strapparava, eds. *Lecture Notes in Computer Science, Adaptive Hypermedia and Adaptive Web-Based Systems, Conference Proceedings AH2000*, Trento, Italy, 267-238, Springer.

Tergan, S., 1998. Checklists for the Evaluation of Educational Software: Critical Review and Prospects. *In: Innovations in Education and Training International,* Vol 35, No.1. pp 9-20.

Thimbleby, H., 1996. Internet, Discourse and Interaction Potential. In: eds. L. K. Yong, L. Herman, Y. K. Leung and J. Moyes, *First Asia Pacific Conference Proceedings. On HCI 1996*, pp3-18.

Totterdell, P. and Cooper, P., 1987. Design and Evaluation of the AID Adaptive Front End to Telecom Gold. *In:* eds. D. Diaper and R. Winder, *BCS Workshop Series Conference Proceedings of the 3rd conference of BCS HCI Specialist People and Computers*, pp 281-295, Cambridge University Press.

Trella, M., Conejo, R. and Guzman, E., 2000. A Web-Based Socratic Tutor for Trees Recognition. *In:* P. Brusilovsky, O. Stock and C. Strapparava, eds. *Lecture Notes in Computer Science, Adaptive Hypermedia and Adaptive Web-Based Systems, Conference Proceedings AH2000*, Trento, Italy, pp 239-249, Springer.

Underwood, J., Dahlberg, A., Fitzpartick, S. and Greenwood, M., 1996. A STILE Project Case Study. *In: Association for Learning Technology Journal,* Vol. 4, No. 2. pp 40-47.

Weber, G. and Specht, M.,1997. User Modelling and Adaptive Navigation Support in WWW-Based Tutoring Systems. *In: Conference Proceeding of the 6th International Conference on UM 1997.*pp 289-300.

Wharton, C., Bradford, J., Jeffries, R. and Franzke, M., 1992. Applying Cognitive Walkthroughs to More Complex User Interfaces: Experiences, Issues and Recommendations. *In:* P Bauersfeld, J. Bennett and G. Lynch, eds. *Human Factors in Computing Systems Conference Proceedings CHI 1992.* pp 381-386, ACM Press, Addison Wesley.

Wiil, 1998. Open Hypermedia Systems, Inteopeability and Standards. *In: Journal of Digital Information (JoDI) Vol. 1, No. 2, January 1998. http:jodi.ecs.soton.ac.uk/Articles/v01/i02/editorial.shtml.*

Wills, G., 2000. Design and Evaluation of Industrial Hypermedia. *In: PhD Thesis*, January 2000, University of Southampton.

Wright, P., 1991. Cognitive Overheads and Prostheses: Some Issues in Evaluating Hypertexts. *In: Conference Proceedings Hypertext 1991.*

Wu, H., De Bra, Aerts, A. and Houben, G., 2000. Adaptation Control in Adaptive Hypermedia Systems. *In:* P. Brusilovsky, O. Stock and C. Strapparava, eds. *Lecture Notes in Computer Science, Adaptive Hypermedia and Adaptive Web-Based Systems, Conference Proceedings AH2000*, Trento, Italy, 250-259, Springer.

Young, R., Howes, A. and Whittington, J., 1990. A Knowledge Analysis of Interactivity. *In:* D. Diaper, eds. *Human Computer Interaction, INTERACT 1990*, pp 115-120, Elsevier Science Publishers B.V.

Zissos, A. and Witten, I. 1985. User Modelling for a Computer Coach: A Case Study. *In: International Journal of Man-machine Studies*, Vol. 23, pp729-750.