Investigation of the Use of Navigation Tools in Web-based Learning: A Data Mining Approach

Christine G. Minetou, Sherry Y. Chen, and Xiaohui Liu

School of Information Systems, Computing and Mathematics

Brunel University, Uxbridge, Middlesex, UB8 3PH, UK

{Christine.Minetou, Sherry.Chen, Xiaohui.Liu}@brunel.ac.uk

Running Head: Mining the Use of Navigation Tools

Requested for reprints should be sent to Sherry Y. Chen, School of Information Systems, Computing and Mathematics, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK, Email: Sherry.Chen@brunel.ac.uk

Abstract

Web-based learning is widespread in educational settings. The popularity of web-based learning is due in great measure to its flexibility. Multiple navigation tools provided some of this flexibility. Different navigation tools offer different functions. Therefore, it is important to understand how the navigation tools are used by learners with different backgrounds, knowledge, and skills. This paper presents two empirical studies in which data mining approaches were used to analyze learners' navigation behavior. The results indicate that prior knowledge and subject content are two potential factors influencing the use of navigation tools. In addition, the lack of appropriate use of navigation tools may adversely influence learning performance. The results have been integrated into a model that can help designers develop web-based learning programs and other web-based applications that can be tailored to learners' needs.

1. Introduction

E-learning is becoming an essential part in educational settings. Current course design and instructional methods are influenced by the development of modern technology (Ngai, Poon, and Chan, 2007). Technological advances in computing and communication have allowed the transition from traditional classrooms to virtual learning spaces. Currently, a growing number of universities worldwide are offering web-based learning programs. In addition, numerous companies are using web-based learning to train their employees. The popularity of web-based learning reflects its flexibilities (Chen and Macredie, 2004).

Unlike traditional computer-based learning, web-based learning programs offer multiple navigation tools for learners to develop their own navigation strategies. Among various navigation tools, alphabetical indices, hierarchical maps, hypertext links and main menus are most commonly used in web-based learning programs. An alphabetical index provides learners with a means to locate particular information without going through a fixed sequence of information. The index lists all concepts in an alphabetical order and does not provide any indication for the relationships between them. A hierarchical map is a graphical representation of important relationships between concepts. With the map, each key concept can be represented as a hierarchy of concepts going from a higher level of abstraction (general concepts) to lower levels (specific concepts, examples). Hypertext links are a mechanism for learners to move quickly to related pages and may reflect either the connectivity of the subject matter or the organization of the content. A main menu in a web-based learning program is a displayed list of choices referring to the concepts (Khalifa and Kwok, 1999).

It seems that these navigation tools provide different functions, and navigating with these tools is complex (Padovani and Lansdale, 2003). In particular, learners who use web-based learning programs are diverse, in terms of their preferences, skills, and needs. In response to this challenge, more attention needs to be directed to see how diverse populations use the navigation tools provided by web-based learning programs. As suggested by Stanney, Mourant, and Kennedy (1998), understanding human factors is useful in identifying the needs of each individual. Therefore, empirical evaluation of learners' navigation behavior becomes paramount, because such evaluation can provide concrete prescriptions for developing web-based learning programs that can accommodate the learners' needs.

In this vein, two empirical studies reported in this paper aim to examine how human factors influence learners' navigation behavior, especially when using navigation tools. Among a variety of individual difference approaches, this study will focus on prior knowledge, because it plays a central role in the setting of goals, which, in turn, determines learners' navigation paths (Corredor, 2004). In addition to navigation behavior, the effects of learning performance are also investigated. By doing so, a sound understanding can be provided not only for which tools learners *prefer to use,* but also for which tools *are most suited* to them. In other words, two research questions were investigated: (a) what are the effects of prior knowledge on the use of navigation tools and (b) what are the effects of the use of navigation tools on learning performance. Answers to these two questions are sought by using a data mining approach to analyze students' navigation behavior because data mining has been successfully applied in many fields to help discover knowledge and to help make decisions, including bioinformatics (Liu and Kellam, 2003), financial analysis (Thawornwong and Enke, 2004), and information retrieval (Zhang and Dong, 2002).

The paper is structured as follows. Section 2 presents research rationale by analyzing the problems of existing work in the field. Section 3 describes the methodology used to conduct two empirical studies and the techniques applied to the analysis of the corresponding data. Subsequently, the grouped navigation behavior of two empirical studies is presented in Section 4, where the effects of prior knowledge on learners' navigation behavior and the relationships between navigation behavior and learning performance are also discussed. Finally, conclusions are drawn and possibilities for future work are identified in Section 5.

2. Research Rationale

Learning is a process of making connections between what is already known and new information (Anderson, Spiro, and Anderson, 1978). Thus, prior knowledge is essential to the learning process. As claimed by Ausubel (1968), the most important single factor influencing learning is what learners already know. Previous research indicates that experts and novices, who have different levels of prior knowledge, demonstrate different learning characteristics. Their different characteristics are summarized in Table 1. In the past decade, research into the influence of prior knowledge on the use of web-based learning programs has mushroomed. A number of studies have found that learners with different levels of prior knowledge benefit differently from the web-based learning programs, with experts and novices requiring different levels of navigational support (McDonald and Stevenson, 1998a, 1998b; Calisir and Gurel, 2003). Chen et al., (2006) present a review of relevant research covering 26 quantitative and qualitative studies. One of the conclusions in their review is that experts are more interested in using tools that could facilitate the location of detailed information related to specific entities (Farrell and Moore, 2000), while novices seem to benefit from hierarchical maps that can facilitate the integration of individual topics (Potelle and Rouet, 2003).

{Table 1 Here}

While this conclusion has provided initial guidance on the role of navigation tools, some problems are still being ignored. One of them is whether subject content also influences learners' navigation behavior. Anderson (1983) divided learning into two states: *declarative* and *procedural* knowledge. Derry (1990) distinguishes between these two, declarative being "knowledge that" and procedural being "knowledge how". Declarative knowledge refers to knowledge about the world and its properties while procedural knowledge refers to knowledge of how to do things (McGilly, 1994). Sun, Merrill, and Peterson (2001) proposed another distinction based on consciousness processing. Learners tend to have conscious access to declarative knowledge but not to procedural knowledge. In other words, declarative knowledge is explicit knowledge whereas procedural knowledge is implicit knowledge (Dienes and Perner, 1999). Chen (2005b) found that the learning performance of obtaining declarative knowledge was influenced by learners' perceptions to the design of learning programs. On the other hand, the learning performance of getting procedural knowledge was affected by their prior knowledge. However, paucity of research examines whether learners' navigation strategies in taking learning programs to present subject content on declarative knowledge are different from those in taking learning programs on procedural knowledge.

Another problem is the majority of previous studies use statistical analyses, which indicate some trends about learners' navigation behavior. However, these are not enough to build learner communities, who demonstrate similar navigation behavior. For such a task, data mining is more appropriate because it can search for valuable information in large volumes of data (Hand et al., 2001). The main difference between statistical analyses and data mining lies in the aim that is sought. The former is used to verify existing knowledge in order to prove a known relationship (Moss and Atre, 2003) while the latter is aimed at finding unexpected relationships (Wang, Rees, and Liao, 2002). As opposed to traditional experiments designed to verify *priori* hypotheses with statistical analyses, data mining uses the data itself to uncover relationships and patterns. By doing so, hidden relationships, patterns, and interdependencies can be discovered and predictive rules can be generated, which are the advantages of data mining (Hedberg, 1995; Gargano and Raggad, 1999).

Much of the work in data mining can be divided into three major categories based on the nature of their information extraction: classification, clustering, and association rules (Chen and Liu, 2004). Clustering, a major *Exploratory Data Analysis* method (Tukey, 1977), is concerned with the division of data into groups of similar objects. Each group, called a cluster, consists of objects that are similar between themselves and dissimilar to objects of other groups (Roussinov and Zhao, 2003). This technique has the advantage of uncovering unanticipated trends, correlations, or patterns. Also, no assumptions are made about the structure of the data. Wang et al. (2004) have developed a recommendation system for the cosmetic business. In the system, they segmented the customers by clustering algorithms to discover different behavior groups. Customers in same group have similar purchase behavior. Classification refers to the data mining problem of attempting to discover predictive patterns where a predicted attribute is nominal or categorical. The predicted attribute is called the class. Subsequently, a data item is assigned to one of a predefined set of classes by examining its attributes (Changchien and Lu, 2001). In other words, the objective of classification is not to explore the data to discover interesting segments, but rather to decide how new items should be classified. For example, Esposito, Licchelli, and Semeraro (2004) built student models for an e-learning system based upon the student performance evaluation: good, sufficient or insufficient. Association rules that were first proposed by Agrawal and Srikant (1994) are mainly used to find out the meaningful relationships between items or features that occur synchronously in databases (Wu, Zhang, and Zhang, 2002). This approach is useful when one has an idea of the different associations being sought. This is because one can find many different correlations in a large data set. Cunningham and Frank (1999) applied the association rules to the task of detecting subject categories that co-occur in transaction records of books borrowed from a university library. As shown by the aforementioned studies, data mining opens a new window for data analyses. We, therefore, used a data mining approach to analyze learners' navigation behavior in two empirical studies conducted in web-based learning programs, which are described in the next section.

3. Methodology Design

3.1 Empirical studies

- ***Web-based Learning Programs***

As described in Section 2, learning can be divided into two conditions: procedural knowledge and declarative knowledge. The two empirical studies described in this paper attempted to identify their impacts on web-based learning. Study One investigated the effects of prior knowledge in a web-based learning program that taught students procedural knowledge, i.e. *How to Use HTML* (Figure 1), and Study Two examined this issue in another web-based learning program that delivered declarative knowledge to the students, i.e. *Computational Algorithms* (Figure 2). The other reason of choosing these two topics was that they were related to students' lectures. As recommended by Reeves (1993), an optimal scenario for answering research questions would be conducted as a study in an environment and context that is meaningful and relevant to the sample population.

{Figures 1, 2 Here}

To identify the learners' different navigation behavior, both web-based learning programs provided them with a variety of navigation tools, including an alphabetical index, a hierarchical map, a main menu and section buttons. In addition, there were rich hypertext links within the text. In this way, the learners were given the freedom to decide their own navigation strategies so that the navigation tools of their choices could be identified by examining their navigation behavior.

- ***Participants***

In order to find reliable evidence about the use of navigation tools, two empirical studies were conducted. 65 learners participated in Study One and 69 learners in Study Two. Participants were undergraduate students in a university in the UK and they volunteered to take part in the studies. A request was issued to students in lectures, and further by email, making clear the nature of the studies and their participation. All participants had the basic computing and Internet skills necessary to operate a web-based learning program.

- ***Procedure***

Both Study One and Study Two consist of four testing steps, which are illustrated in Figure 3. All of the participants interacted with the web-based learning (WBL) programs for about two hours and their interactions were collected by recording them in a log file, including the frequencies of access of navigation tools named hierarchical map, alphabetical index, section buttons, hypertext links and main menu. These data were analyzed by using data mining techniques (See Section 3.2).

{Figure 3 Here}

The questionnaire was applied to measure students' prior knowledge. The questions were classified into two major groups based on their homogeneity: (a) *domain knowledge*: the preliminary understanding of subject content, for example, How familiar are you with designing Web pages with HTML?; (b) *system experience*: the experience of using a variety of systems, for example, How much do you enjoy accessing the Internet?. Each question used a five-point Likert Scale consisting of: very much; quite a lot; average; not much; not at all. The last two were recognized as novices and the remaining three were identified as experts. To reduce the bias of this study, other human factors such as culture background and gender differences were also examined using the questionnaire.

The pre-test and post-test were designed to assess the participants' performance both before and after using the web-based learning programs. Both included 20 multiple-choice questions, each with four different answers and a "don't know" option, from which the students could choose only one. The students' learning performance was measured based on *gain scores*, which was calculated as the post-test score minus the pre-test score. The details of the design rationale of the questionnaire, pre-test, and post-test for Study One can be found in Chen and Macredie (2004) and for Study Two in Mitchell et al. (2005).

3.2 Data analysis

As indicated in Section 2, much of the work in data mining can be divided into clustering, classification, and association rules. Among these three approaches, clustering is selected for analyzing data of the aforementioned two studies because it can form groups that share similar characteristics (Nolan, 2002). More precisely, given a set of data points, each having a set of attribute values, clustering is the process of grouping the data points into different clusters using a dissimilarity measure such that data points in the same cluster are more similar to one another and those in different clusters are less similar to one another. The principle of clustering is to maximize the similarity between all points inside a cluster and minimize the similarity between the different clusters (Han and Kamber, 2001).

There are three different general types of clustering algorithms: those based on the attempt to find the optimal partitions into a specific number of clusters, those based on a hierarchical attempt to discover cluster structures and those based on probabilistic models for underlying clusters (Hand, et al., 2001). The choice of the most appropriate type of algorithm is strongly related to the objectives of the task as well as the nature of the data. The aim of two studies described in this paper is to group users into clusters based on attributes that originate from the use of navigation tools within the web-based learning programs so partition-based clustering is used here for analyzing the data. The most well-known and commonly used partition-based clustering is K-means (Han and Kamber, 2001). However, the challenge of using the K-means algorithm is that the number of clusters needs to be fixed in advance. Therefore, there is also a need to find out the most appropriate number of clusters with probabilistic clustering methods, of which the Expectation Maximization (EM) algorithm is used. In summary, K-means was selected for the formation of the final clusters and EM algorithm was used to gain an insight about the optimal number of clusters. A main disadvantage of the probabilistic approach is the complexity of the associated estimation algorithm. Partition-based clustering methods, such as K-means, on the other hand, have the advantage of being simple and transparent (Hand, et al., 2001). The approaches of using these two algorithms are described below.

- ***EM algorithm***

The EM algorithm is a statistical model that makes use of Gaussian mixtures model. A mixture is a set of N probability distributions where each distribution represents a cluster. EM assigns a probability to each data point. It is the probability that the data point would have, if it were to have a certain set of attribute values, given it was a member of a specific cluster. The probability distribution estimates the membership of each data point to each of the clusters. Consequently, these probabilities are used as the basis of partitioning the data and hence defining the clustering (Alpaydin, 1998).

More specifically, the EM algorithm iterates between two steps for estimating parameters of generative models based on available data. The two steps are the Expectation step and the Maximization step. The Expectation step deals with the unknown underlying variables using the current estimate of the parameters and conditioned upon the observations. The Maximization step then provides a new estimation of the parameters. The two steps are iterated until a desired convergence value is achieved. The convergence value that constitutes the stationary point is the one that maximizes the likelihood, but it is not guaranteed to be the optimal one, as the algorithm can terminate upon reaching a local maximum. To overcome this problem, different initial values or random numbers should be used (Stoica and Selén, 2004).

EM can be viewed as an iterative optimization algorithm for maximizing a likelihood score function given a probabilistic model with missing data. In the present study, the mixture model can be regarded as a distribution in which not the variables but the class labels are missing (Hand, et al.,2001). In contrast to clustering algorithms such as hierarchical clustering and K-means clustering, EM can generate clusters based on cross validation without requiring any input on the number of clusters. As a result, it is possible to find an indication for the 'optimal' number of clusters in a specific dataset by exploiting the algorithm (Chen, 2005a). The EM algorithm was applied to both datasets from Study One and Study Two. For both datasets, the EM algorithm groups learners into three clusters for different assignments of initial random numbers.

- ***K-means algorithm***

A re-estimation procedure was conducted in K-means algorithms in order to cluster a given data set through a certain number (*N*) of *K* clusters fixed a priori. In the first step, the core idea is to define *K* centroids (cluster centers), for each cluster. Another parameter called seed (*S*) is used to generate the random numbers for the assignment of the initial centroids. Since K-means is sensitive to how clusters are initially assigned (Pena, et al, 1999), it is necessary to try different random values and evaluate the results in order to find which combination fit better to the data. This can be achieved by varying the value of *S*. In detail, the algorithm consists of the following: the first step is to define *K* centroids or average of the points, one for each cluster. Following that, each point is associated to the nearest centroid. The next step starts when all the points are assigned to clusters; it is the re-calculation procedure of *K* new centroids. For these *K* new centroids, a new binding has to be done between the same data set points and the nearest new centroid. These two steps are alternated until a stopping criterion, named also cost function (Sun and Wang, 2001), is met, i.e., when there is no further change in the assignment of the data points. K-means uses the squared-error cost function, which is an indicator of the distance of the given points from their respective cluster centers; the aim is to minimize distance. A mentioned before, in this particular case, K-means uses the Euclidean distance measure to compute distances between a data point and a cluster center. The K-means performance reveals the centroid of each cluster as well as statistics on the number and percentage of instances assigned to different clusters. Thus, centroids can be used to characterize each one of the formed clusters.

Despite of the insight from the EM algorithm about the number of clusters, different combinations in the number of clusters and seed values have been tested in order to reveal which combination provides the best algorithm performance and to check whether the outcome is consistent with what the EM algorithm recommends. The performance of the algorithm was evaluated according to: (a) the value of the sum of squared errors; (b) the percentage of clustered instances in each cluster; (c) the mean value of each attribute within the cluster and (d) their visual representation (data not shown). Results are in accordance with the EM algorithm, which indicated that the K-means algorithm performs better when the number of the pre-defined clusters are N=3 and S=20 for Study One and N=3, S= 10 for Study Two. The sum of squared errors within clusters for Study One is 16.9 and the one for Study Two is 22.1.

4. Results and Discussions

There are five navigation tools and the descriptive statistics for each navigation tool is showed in Table 2. To correspond with human factors, the results of both studies have indicated that prior knowledge of subject domain has a great influence on the use of navigation tools (Section 4.1). In addition, we found that the use of navigation tools has an impact on the students' learning performance (Section 4.2).

{Table 2 Here}

4.1 The Use of Navigation Tools

{Table 3 Here}

4.1.1 Study One

The left column of Table 3 reveals the attributes that characterize each cluster in Study One. The percentage of learners within each cluster is satisfactory for the total number of 65 instances (learners). Clusters can be characterized as roughly balanced: Cluster 1 (N = 18): 28%, Cluster 2 (N= 25): 38%, Cluster 3 (N = 22): 34%. The mean values of the above attributes, i.e., the average frequencies each navigation tool (including main menu, hierarchical map, alphabetical index etc) is chosen by learners, are shown in the table, which has indicated that the learners are grouped according to the following trends:

- Cluster 1 (C1): learners frequently used the hypertext links and seldom used the main menu, hierarchical maps, alphabetical index, and section buttons;
- Cluster 2 (C2): learners frequently used the alphabetical index, occasionally used the hypertext links, and seldom used main menu, hierarchical map, and section buttons;
- Cluster 3 (C3): learners frequently used the hierarchical maps, occasionally used the hypertext links, and seldom used the main menu, section buttons, and alphabetical index.

{Figure 4 Here}

To find the corresponding human factors for each cluster, the results indicated that learners with different levels of prior domain knowledge, i.e. the preliminary understanding of HTML, appear in different clusters. The experts who had a high level of HTML knowledge mainly emerge in Cluster 2, in which the learners visited the alphabetical index many more times (Figure 4). The alphabetical index is useful for locating specific information (Chen and Macredie, 2002). This finding is in line with that of the study by Carmel *et al* (1992), which found that high-knowledge users were more interested in using tools that could facilitate the location of detailed information related to specific entities. Conversely, the majority of novices who had a low level of HTML knowledge appear in Cluster 1 or Cluster 3, in which learners often used the hypertext links or hierarchical map respectively. A possible explanation for this finding is that the mental structure of novices is more chaotic and disorganized (Spires and Donley, 1998). Therefore, they need to rely on the hypertext links to connect the relevant concepts or to use hierarchical map to identify the content structure. In particular, the hierarchical map not only reveals the *document structure*, i.e. the physical arrangement of a document, but also reflects the *conceptual structure*, i.e. the relationships between different concepts (Nilsson and Mayer, 2002). In other words, the hierarchical map can help novices incorporate the document structure into the conceptual structure, which can facilitate the integration of individual topics (Dee-Lucas, and Larkin, 1995; Möller and Müller-Kalthoff, 2000).

4.1.2 Study Two

The navigation behavior of the three clusters in Study Two is presented in the right column of Table 3. The mean value of the above attributes indicates that learners are grouped according to the following trends:

- Cluster 1 (C1): learners occasionally used the hypertext links and seldom used any other navigation tools;
- Cluster 2 (C2): learners frequently used the hypertext links, occasionally used section buttons, and seldom used the main menu, hierarchical maps, and the alphabetical index;
- Cluster 3 (C3): learners frequently used the hypertext links, occasionally used section buttons and hierarchical maps, and seldom used the main menu and alphabetical index.

According to the percentage of students within each cluster, clusters can be considered as very unbalanced for the total number of 69 instances (learners): Cluster 1 (N = 40): 58%, Cluster 2 (N= 24): 35%, Cluster 3 (N = 5): 7%. In other words, the majority of learners appear in Cluster 1, in which the hypertext links were occasionally used and other navigation tools were rarely used by the students. This is very different from the findings of Study One. It may be due to the fact that the web-based learning programs of these two studies present different subject content. As indicted in Section 3, the subject content in Study One presents procedural knowledge of how to use HTML, while Study Two introduces declarative knowledge about the properties of computation algorithms. Procedural knowledge is not available to awareness (Dienes and Perner, 1999). Thus, the learners can appreciate the advantages of using navigation tools when they learn procedural knowledge. Conversely, the learners are aware of declarative knowledge in the existing content. Therefore, they may not consider navigation tools as useful mechanisms when the subject topic is related to declarative knowledge. This issue suggests that subject content is another critical factor that should be considered in the design of navigation tools for web-based learning programs.

Like Study One, the learners' prior domain knowledge, i.e. their preliminary understanding of computational algorithms, is a dominant human factor for each cluster. As showed in Figure 5, all learners in Cluster 3 are novices who had a low level of knowledge about computational algorithms. The main difference between Cluster 3 and other clusters is that the hierarchical map is a preferred tool in Cluster 3. This is consistent with the findings of Study One, which indicate that the hierarchical map is favored by the novices because it can help them to build an integrated picture of the subject content. On the other hand, it may not be useful for experts to locate the details of specific items.

{Figure 5 Here}

4.1.3 Overall Use

The overall use of navigation tools for Study One is presented in Figure 6 and Study Two in Figure 7. It seems clear that the hypertext links are the most favored tool in both Study One and Study Two. The web-based learning programs employ hypermedia techniques (Federico, 2000), of which the major feature is the capability to link relevant pages (Bar-Ilan, 2005). The aforementioned finding confirms the value of this feature, which encourages learners to navigate by association. In this way, they can make learning paths in their own self-directed manner, instead of having to follow passively some form of pre-defined linear access (Farrell and Moore, 2000).

Section buttons are the second favored tool in Study Two, while they are the least used tools in Study One. A possible reason is that the subject content of Study Two has a specific division and includes 6 sections: (1) Introduction to Algorithms, (2) Asymptotically Slow Sorts, (3) Divide and Conquer, (4) Analysis of Algorithms, (5) Searching Algorithms, and (6) Background Mathematics. Each section is further split into four sub-topics. On the other hand, the subject content of Study One is roughly divided into 3 sections: (1) What is HTML?; (2) Working with HTML; and (3) Relations with SGML and WWW. Section 2 is the key element of the web-based instructional program, which covers 12 sub-topics of HTML authoring. In other words, the content structures of these two studies are different, which suggest that the content structure of the web-based learning programs also influences the choice of navigation tools.

{Figures 6, 7 Here}

4.2 Impacts on Learning Performance

The results of Section 4.1 indicate that the learners of different clusters have different levels of prior knowledge. Prior knowledge seems to influence the learners' behavior. To investigate whether the learners' behavior influences their performance, the learning performance of each cluster showed in Table 4 was analyzed on the basis of the *gain score*.

{Table 4 Here}

4.2.1 Study One

The left column of Table 4 reveals the learning performance of each cluster in Study One. Experts who were in Cluster 3 had a better performance than those who were in other clusters. On the other hand, novices in Cluster 2 performed slightly better than those in Clusters 1 and 3. This is not consistent with the findings of the navigation behavior presented in Section 4.1.1, which indicates that most of the experts appeared in Cluster 2, where the alphabetical index was frequently used. Conversely, the novices mainly emerged in Cluster 1 or Cluster 3, where the hypertext links and hierarchical map were often selected. This implies that experts may prefer to use the alphabetical index although it may not be beneficial for their learning performance. In contrast, the novices may prefer to select the hypertext links and hierarchical map but these two tools may not be helpful for improving their learning performance. These findings imply that what learners like may not be what they need. The other explanation is that preferences and performance are two different things. Preference is defined by a function that represents how much a user likes or dislikes a given item (Jung, Hong, and Kim, 2005) while performance is the ability of the learners to actually solve problems (Topi and Lucas, 2005). Competent performance requires not only perquisite knowledge and skills but also beliefs of personal efficacy to use both effectively (Mavis, 2001). The findings of this study suggest that there is no direct relationship between performance and preferences. These results do not echo the findings of previous studies (Ford and Chen, 2001; Fullerton, 2000), which suggest that matching the design of learning programmes with learners' preferences can enhance their performance.

4.2.2 Study Two

The learning performance of three clusters in Study Two is presented in the right column of Table 4. It seems that the learners who were in Cluster 2 performed best and those in Cluster 1 did worst, especially the novices. As indicated in Section 4.1.2, the learners in Cluster 2 frequently clicked the hypertext links and those in Cluster 1 rarely used the navigation tools, apart from choosing the hypertext links occasionally. This implies that the absence of the appropriate use of navigation tools might hinder the learners' performance, which is in line with the finding of the study by Tung el al. (2003). As indicated by Jul and Furnas (1997), navigation is the process of moving around an environment, deciding at each step where to go. Navigation tools help to support the learner's navigation through the site, performing two functions: telling the learner what information is held within the site; and helping them to find the information quickly and easily (Park and Kim, 2000). However, the majority of learners appeared in Cluster 1. This indicates that the learners did not appreciate the value of the navigation tools even though using them would have been beneficial. These findings reiterate those of Study One that the learners are probably not aware of what they need in their learning process. Therefore, the designers of web-based learning programs should not only consider accommodating learners' different preferences, but also should investigate how to provide support for learners to identify suitable navigation tools.

5. Development of a Model

The results presented in the aforementioned section are very interesting. Based on the key results, Figure 8 presents a model that illustrates the impact of prior knowledge and subject content on the use of navigation tools and the effects of the use of navigation tools on learning performance. This model can help designers understand more easily "what" learners' need and "why" they need. The designers can then develop web-based programs and other web-based applications that can tailor navigation support to the learners' knowledge, skills, and tasks.

{Figure 8 Here}

- ***Learners with different levels of prior knowledge prefer to use different navigation tools.***

The results of these two studies reveal that learners with different levels of prior knowledge benefit from different navigation tools. The hierarchical map was favored by novices while the alphabetical index was preferred by experts. It may be due to the fact that novices lack the prior knowledge and the hierarchical map presents the content in a structured format and can thus help novices organize content (Calisir and Gurel, 2003). On the other hand, experts have acquired a great deal of prior knowledge and they are more able to impose structure on the content (Spires and Donley, 1998). Therefore, they tend to use navigation tools that can provide them with free navigation and help them find specific information. The alphabetical index is one type of such tools.

- ***Subject content is a potential issue that influences the use of navigation tools.***

The web-based learning programs used in the two studies introduced different subject contents. Their subject contents represent different types of knowledge. The web-based learning program used in Study One was related to procedural knowledge and that in Study Two was concerned with declarative knowledge. As indicated in Section 2, this difference can be equated to the distinction between implicit and explicit knowledge, because procedural knowledge is generally inaccessible to consciousness while declarative knowledge is accessible (Sun and Zhang, 2004; Sun, Merrill and Peterson 2001). When learners interact with subject content about procedural knowledge (implicit knowledge), they may need to rely on visual cues to access information. Navigation tools offer such visual cues. This may be the reason why learners in Study One used the navigation tools more frequently than those in Study Two.

- ***The use of navigation tools may have effects on learners' performance.***

Web-base learning employs hypermedia techniques to present the content in a non-linear format and to provide learners with great freedom to sequence the information. However, this freedom may cause disorientation problems (Nielsen, 2001). Navigation tools provide learners with additional visual cues, which can help them structure the content and reduce the disorientation problems. As suggested by de Jong and van der Hulst (2002), visual cues provide learners with a systematic route through the domain and may thus lead to a better acquisition of the structure of the domain. Learners do not, however, take advantage of such visual cues provided by navigation tools so the acquisition of their knowledge is influenced. This may be able to explain the results of Study Two that the lack of the appropriate use of navigation tools may obstruct learners' performance.

6. Conclusions

This paper presents two empirical studies, in which data mining approaches were applied to answer two research questions. In response to the first research question, *what are the effects of prior knowledge on the use of navigation tools*, we have found that the students' prior knowledge plays an influential role in their use of navigation tools. In terms of the second research question, *what are the effects of the use of navigation tools on learning performance*, we have demonstrated that the use of navigation tools has great impact on the students' learning performance. The results also suggest that the degree of the use of navigation tools is influenced by subject topics. In brief, prior knowledge, subject content, and learning performance are three issues that should be considered in the design of navigation tools for web-based learning programs.

These two empirical studies described in this paper have shown the importance of understanding the use of navigation tools in web-based learning. However, they were only small-scale studies. Further work needs to be undertaken with a larger sample to provide additional evidence. Another limitation of these studies is that the sample was not very balanced. There were fewer experts and more novices in Study One, while more experts and fewer novices in Study Two. This limitation may influence the validity of the results in that the standard deviation is generally high. Further empirical studies are needed to verify the results described in this paper.

Given any dataset, there are often no strict rules that impose the use of one specific method over another in its analysis. Therefore, there is a need to analyze learners' navigation behavior using other clustering algorithms or even other data mining approaches, e.g. classification and association rules. It would be interesting to see what results will be found by using these methods. Gathering information on these issues through further work can help clarify the findings from the present study. In addition, the results of such studies could be integrated to build robust user models for the development of effective web-based learning programs that can accommodate the need of each individual learner.

References

Agrawal, R. & Srikant, R. (1994). Fast algorithms for mining association rules. *Proceedings of the 20th international conference on very large databases.* Santiago, Chile.

Alpaydın, E. (1998). Soft vector quantization and the EM algorithm, *Neural Networks,* 11(3), 467-477.

Anderson, J.R. (1983) The Architecture of Cognition. Cambridge, MA: Harvard University Press.

Anderson, R.C., Spiro, R.J., & Anderson, M.C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal, 15*, 433-440.

Ausubel, D. P. (1968). Educational Psychology: A cognitive view. New York: Holt, Rinehart and Winston.

Bar-Ilan, J. (2005). What do we know about links and linking? A framework for studying links in academic environments. *Information Processing & Management*, 41(4), 973-986.

Calisir F. & Gurel Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior* 19, 135-145.

Carmel, E., Crawford, S., & Chen, H. (1992). Browsing in Hypertext: A cognitive study. *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 865-884.

Changchien, S. & Lu, T. (2001). Mining association rules procedure to support on-line recommendation by customers and products fragmentation. *Expert Systems with Applications*. 20(4), 325-335

Chen, C. (2005a). Measuring the movement of a research paradigm. Proceeding of SPIE-IS&T: Visualization and Data Analysis 2005 (VDA 2005), pp. 63-76. San Jose, CA.

Chen, S. Y. (2005b). Evaluating the Learning Effectiveness of using Web-based Instruction: An Individual Differences Approach. *International Journal of Information & Communication Technology Education.* 1(1), 69-82.

Chen, S. Y. & Liu, X. (2004). The Contribution of Data Mining to Information Science. *Journal of Information Science*. 30(6), 550-558

Chen, S. Y. & Macredie, R. D. (2002). Cognitive Styles and Hypermedia Navigation: Development of A Learning Model. *Journal of the American Society for Information Science and Technology*. 53(1), 3-15.

Chen, S. Y. & Macredie, R. D. (2004). Cognitive Modelling of Student Learning in Web-based Instructional Programmes. *International Journal of Human-Computer Interaction*. 17(3), 375-402

Chen, S. Y., Fan, J, & Macredie, R. D. (2006). Navigation in Hypermedia Learning Systems: Experts vs. Novices. *Computers in Human Behavior*. *Computers in Human Behavior*. 22(2), 251-266

Corredor, J. (2006). General and domain-specific influence of prior knowledge on setting of goals and content use in museum websites. *Computers & Education*, 47(2), 207-221.

Cunningham, S. & Frank, E. (1999). Market Basket Analysis of Library Circulation Data. *Proceedings of the Sixth International Conference on Neural Information Processing*, Volume II, pp. 825-830.

Dee-Lucas, D., & Larkin, J.H., (1995). Learning from electronic texts: effects of interactive overviews for information access. *Cognition and Instruction*. 13(3), 431-468

Derry, S. J. (1990). Learning strategies for acquiring useful knowledge. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 347-379). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dienes, Z. & Perner, J. (1999). A theory of implicit and explicit knowledge. Behavioral and Brain Sciences. 22(5), 735-755.

Esposito , F., Licchelli, O., & Semeraro, G. (2004). Discovering Student Models in e-learning Systems. *Journal of Universal Computer Science*, 10(1), 47-57

Farrell, I.H., & Moore, D.M. (2000). The effects of navigation tools on learners' achievement and attitude in a hypermedia environment. *Journal of Educational Technology Systems*, 29 (2) 169-181.

Federico, P. (2000). Learning styles and student attitudes toward various aspects of network-based instruction. *Computers in Human Behavior*, 16(4), 359-379.

Ford, N. & Chen, S. Y. (2001). Matching/Mismatching Revisited: An Empirical Study of Learning and Teaching Styles. *British Journal of Educational Technology*, 32 (1) 5-22.

Fullerton, K. (2000). *The interactive effects of field dependence-independence and Internet document manipulation style on student achievement from computer-based Instruction*. Ed.D Dissertation. University of Pittsburgh.

Gargano, M. L. & Ragged, B. G. (1999). Data mining – a powerful information creating tool. *OCLC systems services*. 15(2), 81-90.

Han, J. & Kamber, M. (2001). *Data mining, concepts and techniques: Cluster analysis,* Morgan Kaufmann, San Francisco.

Hand D.J., Mannila H., & Smyth P. (2001). *Principles of data mining*, MIT Press.

Hedberg, S. R. (1995). The Data Gold Rush. Byte, 20(10), 83-88.

Jul, S. & Furnas, G.W. (1997). Navigation in electronic worlds: a CHI'97 workshop. SIGCHI Bulletin 29(4), 44–49.

Jung, S. Y., Hong, J. H., & Kim, T. S. (2005). Preference is defined by a function that represents how much a user likes or dislikes a given *item. IEEE Transactions on Knowledge and Data Engineering.* 17(6), 834-843

Khalifa, M. & Kwok, R. (1999). Remote learning technologies: effectiveness of hypertext and GSS. *Decision Support Systems*. 26(3), 195-207.

Liu, X. & Kellam, P. (2003). Mining Gene Expression Data. In: *Bioinformatics: Genes, Proteins and Computers*. C Orengo, D Jones and J Thornton (eds). p. 229-244.

Mavis, B. (2001). Self-Efficacy and OSCE Performance Among Second Year Medical Students. *Advances in Health Sciences Education*, 6(2), 93-102

McDonald S. & Stevenson R.J. (1998a). The effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors*, 40, 18 –27.

McDonald S. & Stevenson R.J. (1998b). Navigation in hyperspace: an evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers*, 10, 129-142.

McGilly, K. (1994). *Classroom Lesson: Integrating Cognitive Theory and Classroom Practice*. Cambridge, MA: MIT.

Mitchell, T. J. F., Chen, S. Y. & Macredie, R. D. (2005). Hypermedia Learning and Prior Knowledge: Domain Expertise vs. System Expertise. *Journal of Computer Assisted Learning*. 21(1), 53-64.

Möller, J. & Müller-Kalthoff, T, (2000). Learning with hypertext: the impact of navigational aids and prior knowledge. *German Journal of Educational Psychology*. 14 (2/3) 116-123.

Moss, L. & Atre, S. (2003). Business Intelligence Roadmap: The Complete Lifecycle for Decision Support Applications. Addison-Wesley.

Ngai, E.W.T., Poon, J.K.L., & Chan, Y.H.C. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers & Education*. 48(2), 250-267.

Nielsen, J. (2001). Did Poor Usability Kill E-Commerce? http://www.useit.com

Nilsson, R. M. & Mayer, R. E. (2002). The effects of graphic organizers giving cues to the structure of a hypermedia document on users' navigation strategies and performance. *International Journal of Human-Computer Studies*, 57(1), 1-26.

Nolan, J. R. (2002). Computer systems that learn: an empirical study of the effect of noise on the performance of three classification methods, *Expert Systems with Applications,* 23(1), 39-47.

Padovani, S. & Lansdale, M. (2003). Balancing search and retrieval in hypertext: context-specific trade-offs in navigational tool use. *International Journal of Human-Computer Studies*, 58(1), 125-149

Park, J. & Kim, J. (2000). Contextual navigation aids for two World Wide Web systems. *International Journal of Human-Computer Interaction*, 12(2), 193-217.

Pena, J. M., Lozano, P., & Larranaga, P. (1999). An empirical comparison of four initialization methods for the k-means algorithm. *Pattern Recognition Letters,* 20, 1027-1040.

Potelle, H. & Rouet, J. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. International Journal of Human-Computer Studies, 58(3), 327-345

Reeves, T. C. (1993). Pseudoscience in computer-based instruction: the case of learner control. *Journal of Computer-based Instruction*. 20(2), 39-46.

Roussinov, D. & Zhao, J. L. (2003). Automatic discovery of similarity relationships through Web mining. *Decision Support Systems*. 35(1), 149-166

Spires, H. A. & Donley, J. (1998). Prior knowledge activation: Inducing engagement with informational texts. *Journal of Educational Psychology*, 90(2), 249-260.

Stanney, K.M., Mourant, R.R., & Kennedy, R.S. (1998). Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence: Teleoperators and Virtual Environments*, 7(4), 327-351.

Stoica, P. & Selén, Y. (2004). Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher," *IEEE Signal Processing Magazine*, 21, 112-114.

Sun, R. & Zhang, X. (2004). Top-down versus bottom-up learning in cognitive skill acquisition. Cognitive Systems Research, 5(1), 63-89

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: a bottom-up model of skill learning. Cognitive Science, 25(2), 203-244

Sun, S. & Wang, Q. (2001). A new validation index for determining the number of clusters in a data set, *Proceedings of IJCNN -01*, 1852–1857.

Thawornwong, S. & Enke, D. (2004). The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing*, 56, 205-232

Topi, H. & Lucas, W. (2005). Mix and match: combining terms and operators for successful Web searches, *Information Processing & Management,* 41(4), 801-817

Tukey, J. W. (1977) Exploratory data analysis. Reading, MA: Addison-Wesley Publishing Company.

Tung, L. L., Debreceny. R., Chan, Y. G., Chan, A. T. L., & Le, S. E. (2003). Interacting with hypertext: an experimental investigation of navigation tools. *Electronic Commerce Research and Applications*, 2(1), 61-72.

Wang, M., Rees, S. J., & Liao, S. Y. (2002). Building an online purchasing behavior analytical system with neural network. *DataMining III*. (Zanasi, Brebbia and Melli, Eds) WIT Press.

Wang, Y., Chuang, Y., Hsu, M., & Keh, H. (2004). A personalized recommender system for the cosmetic business. *Expert Systems with Applications*, 26(3), 427-434

Wu, X., Zhang, C., & Zhang, S. (2002). Mining Both Positive and Negative Association Rules. *Proceedings of the 19th International Conference on Machine Learning*, pp. 658-665.

Zhang, D. & Dong, Y. (2002). A Novel Web Usage Mining Approach for Search Engine. *Computer Networks*. 39(3), 303-310

Table 1: Learning Characteristics of Experts and Novice (Adapted from Chen, et al, 2006)

Experts	Novices
Global mental models	Local mental models
Directed search	Undirected search (trial and error)
Deep structures	Surface features
Mental simulation of integrated functions and whole application	Mental simulation of isolated functions
Complete analysis deferring details	Incomplete analysis
Depth-first strategies	Breadth-first strategies
Design whole and add pieces	Design pieces
Integrated whole throughout the process	Failure to integrate pieces into a whole …
Find the best solution	Find a (any) solution

Table 2: The Descriptive Statistics for Each Navigation Tool

Navigation Tools	Study One		Study Two	
	Mean	SD	Mean	SD
Main Menu	4.4	2.8	1.6	0.8
Hierarchical Map	5.8	1.6	3.2	0.9
Alphabetical Index	6.0	2.2	1.8	0.4
Hypertext Links	8.9	2.9	36.8	9.1
Section Buttons	3.7	1.2	6.5	2.4

(SD: Standard Deviation)

Table 3: Clusters of Study One and Study Two

Navigation Tools		Study One			Study Two		
		Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
Main Menu	Mean	3.6	3.2	6.4	0.5	1.8	2.6
	SD	2.8	2.7	3.2	0.9	0.7	1.1
Hierarchical Map	Mean	3.8	3.0	10.6	0.4	0.9	8.2
	SD	1.2	1.6	3.0	0.7	0.9	1.6
Alphabetical Index	Mean	3.3	9.6	5.2	2.4	1.7	1.4
	SD	1.7	2.2	2.6	0.4	0.5	0.6
Hypertext links	Mean	10.3	7.9	8.4	20.5	57.0	33.0
	SD	3.1	2.7	3.0	9.0	16.0	12.0
Section Buttons	Mean	2.8	2.2	6.1	5.4	5.6	8.6
	SD	1.2	1.2	1.5	2.5	2.3	3.0
Clustered Instances		18	25	22	40	24	5
Total number of Instances		65			69		

(SD: Standard Deviation)

Table 4: Learning Performance of Study One and Study Two

Learning Performance		**Study One**			**Study Two**		
		Cluster 1	*Cluster 2*	*Cluster 3*	*Cluster 1*	*Cluster 2*	*Cluster 3*
Overall	Mean	8.06	7.96	8.09	0.43	1.79	1.00
	SD	1.26	1.59	0.88	2.72	2.75	3.31
Experts	Mean	7.50	7.50	8.25	1.06	1.50	N/A
	SD	3.50	2.25	1.50	3.22	2.97	N/A
Novices	Mean	8.10	8.28	8.10	-0.10	2.09	1.00
	SD	0.90	0.73	0.72	2.18	2.60	3.31

(SD: Standard Deviation)

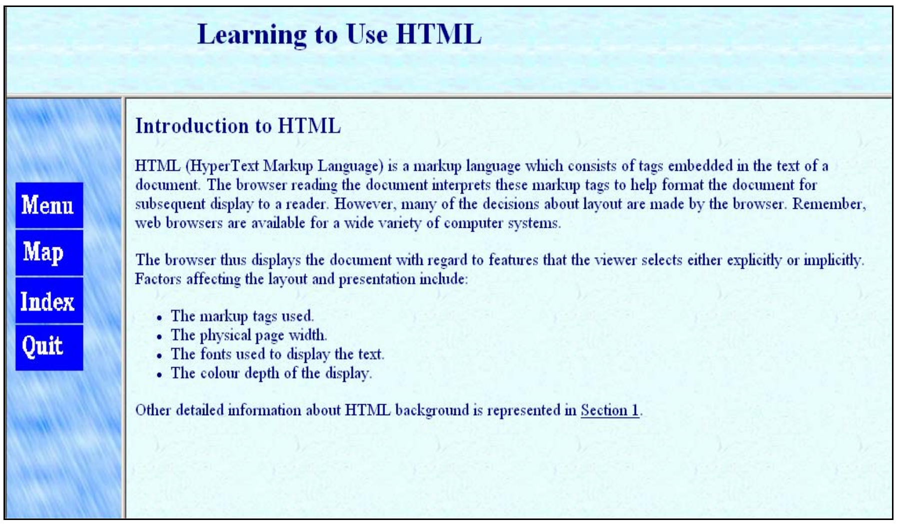

Figure 1: The web-based learning program of Study One

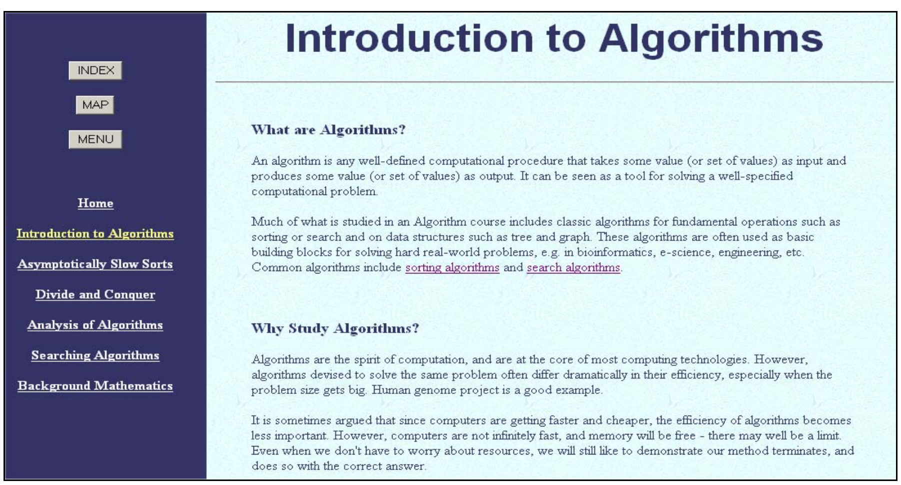

Figure 2: The web-based learning program of Study Two

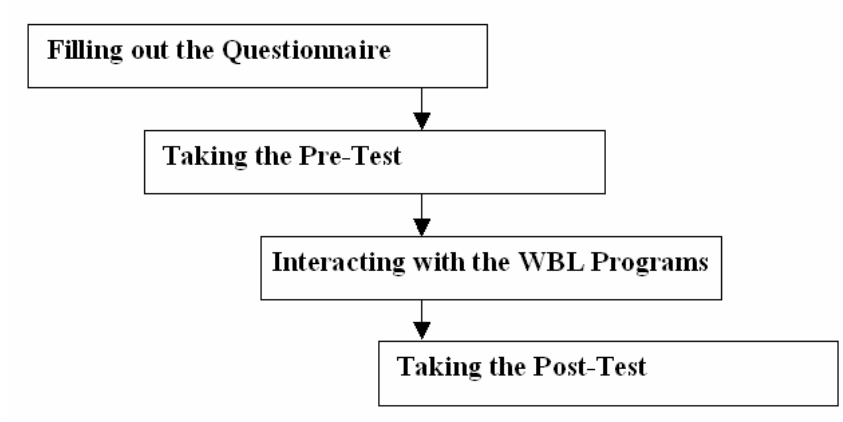

Figure 3: Procedures

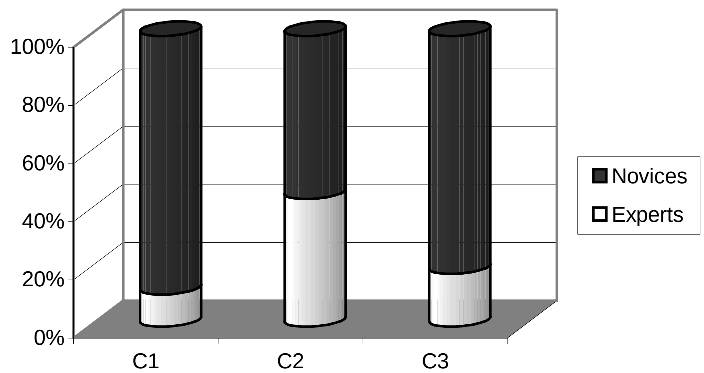

Figure 4: Novices and Experts in Study One

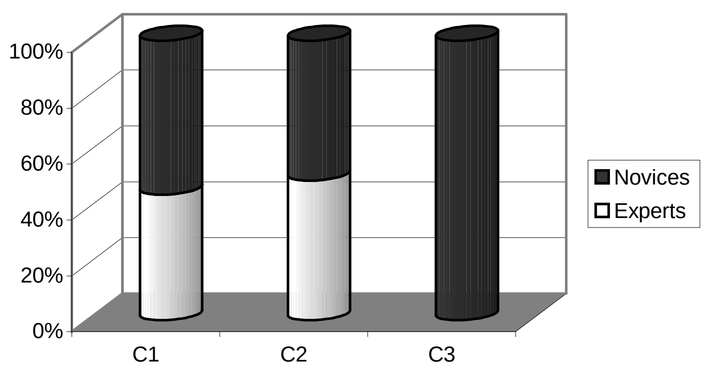

Figure 5: Novices and Experts in Study Two

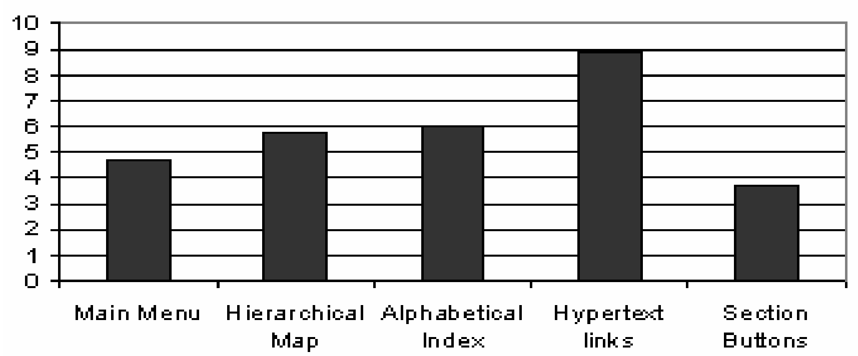

Figure 6: The use of navigation tools in Study One

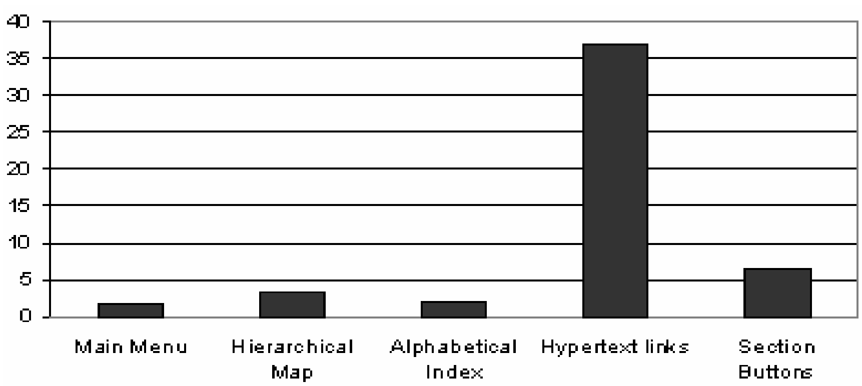

Figure 7: The use of navigation tools in Study Two

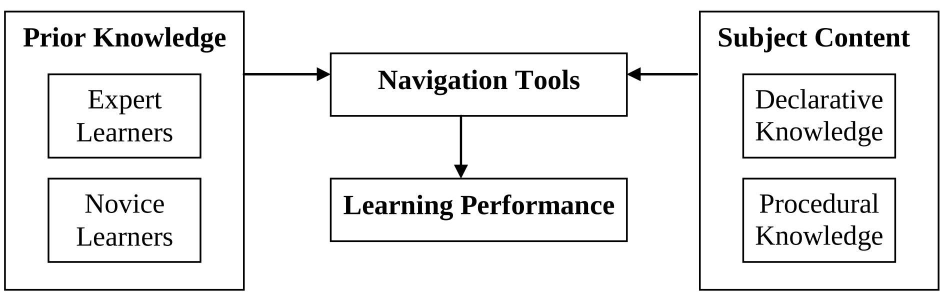

Figure 8: A Model based on the key results