The Influence of Hypertext Linking Structures and Task-Related Variables on Information Retrieval Tasks

by

Michael F. Mohageg

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Industrial and Systems Engineering

APPROVED:

Harry L. Snyder, Chairman

Robert J. Beaton

Paul T. Kemmerling

Michael E. Maddox

William H. Muto

December, 1990

Blacksburg, Virginia

The Influence of Hypertext Linking Structures and Task-Related Variables on Information Retrieval Tasks

by

Michael F. Mohageg

Committee Chair: Harry L. Snyder, Ph.D.

Industrial and Systems Engineering

(ABSTRACT)

Hypertext is a method of online information management and/or presentation where textual documents are parsed (modularized) into many nodes and inter-connected using machine-supported links. These systems have become increasingly popular in numerous applications. Unfortunately, few empirical investigations have been conducted concerning the usability and utility of hypertext, and the effusive claims made by many hypertext enthusiasts are largely unsubstantiated.

This study investigated several usability issues relating to hypertext within the context of an information retrieval application. Of particular interest were system linking structures consisting of linear, hierarchical, network, and combination hierarchical/network configurations. These commonly used hypertext linking structures were imposed on a text-intensive geography database (GEO). GEO contains 187 nodes discussing a variety of topics concerning the countries of North Africa. In addition to the linking structures, the

task variables of number of required links (to reach the answer) and task type were studied. Task type refers to expert programmers' judgements as to whether a task is best suited to a hierarchical or network linking structure.

The approach was to create a set of information retrieval (IR) tasks with specific characteristics (as determined by number of required links and task type), and to study the performance of each linking structure in completing these tasks. The intention was to identify the task situations under which each linking structure excels.

Results indicate that hierarchical linking structures perform quite well for most IR tasks and perform significantly better than network linking. The combination condition performs no worse than hierarchical, yet, with the exception of task completion times, provides no consistent advantages over the hierarchical structure. Hence, it is concluded that, for novice users of a system, the performance advantages resulting from the inclusion of network links (in isolation or in combination with hierarchical) are not commensurate with the associated costs of creating such links. Ultimately, results are aimed at a better understanding of hypertext systems, their performance, and more judicious applications of these systems.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Robert J. Beaton, Prof. Paul T. Kemmerling, and Dr. Michael E. Maddox for serving on the doctoral committee. Their time, helpful comments, and interest are most appreciated. A special thanks is extended to Dr. William H. Muto for his patience, insight, and effort in maintaining my focus on the applied problems of this research. A heartfelt expression of thanks and gratitude is due Dr. Harry L. Snyder. His guidance, competence, and ability to shed light on seemingly hopeless problems have been deeply appreciated in the past four and a half years. Dr. Snyder's trust in me and his providing an environment for autonomous learning, within limits, and independent thought have been instrumental.

I would like to thank my wife, Annick, whose loving support and friendship provided me with constant motivation. Her empathy and unwavering trust were most appreciated. I wish to thank my parents (Rass and Etti Mohageg), siblings, and uncle and aunt (Hass and Dixie Mohaghegh) whose belief in my capabilities propelled me through both undergraduate and graduate work.

This research was sponsored by a graduate internship grant from the Texas Instruments, Inc., User Systems Engineering Laboratory to the Displays and Controls Laboratory. TI's support is appreciated and acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Hypertext and hypermedia have been viewed as the future environments for creating, organizing, and disseminating online documents. In recent years, a great deal of attention and discussion has been devoted to these online environments, especially since the arrival of Apple Computer's HyperCard™ software package, which provided users of personal computers with a significant portion of the capabilities found in high-end hypertext systems. Most of the discussion concerning these systems has been of an effusive and complimentary nature. Authors consistently ponder the potential applications for these systems and repeatedly consider the host of benefits hypertext and hypermedia offer. However, with two exceptions (Cecala, Watkins, and Mohageg, 1989; Gordon, Gustavel, Moore, and Hankey,1988), no solid empirical evidence exists to support or refute the claims and/or predictions made of the basic concepts in the utility and usability of hypertext and hypermedia. While many have considered the potential benefits of these systems and evaluating methods for improving navigation (Conklin, 1987; Jones, 1987), appropriate conceptual models (Vail, 1989), and even the management of hypertext design projects (Shneiderman, 1988), it is reasonable to ask a more fundamental question: is the basic concept of hypertext (and/or hypermedia) appropriate for typical task performance? From an engineering standpoint, satisfaction and efficiency of task performance should be regarded as indicators of usable systems, as opposed to number of different media supported by the system or the number of different features the system provides. Obviously, efficiency and satisfaction will depend on the type of task being performed. The study proposed in this document will attempt to answer several questions concerning the utility and usability of hypertext environments.

This introductory section provides a description and definition of hypertext and hypermedia systems and delineates the difference(s) between these hyper-"environments" and many popular (and more conventional) online information management software systems. The following sections present the historical underpinnings of hypertext and hypermedia and discuss the criticisms made of traditional online systems.

Hypertext and Hypermedia Defined

Hypertext and hypermedia are approaches to computerized information organization enabled for computers by recent technological advancements in computer technologies. "Hyper-environments" have eluded strict definitions since they seem to lack defining characteristics that are either necessary or sufficient. However, the general description provided herein is commonly accepted among professionals working in the area. Hypertext is a method of online information management and/or presentation where textual documents are parsed (modularized) into many nodes. Usually (but not always) each node contains a single concept, data element, idea, or chunk of information. The nodes are connected to one another using machine-supported links (Figure 1); in effect, they create an inter-connected network of nodes. Under idealistic conditions, links can be established among any number of conceptually or semantically related nodes, providing users with the power to access any node by traversing the appropriate link.

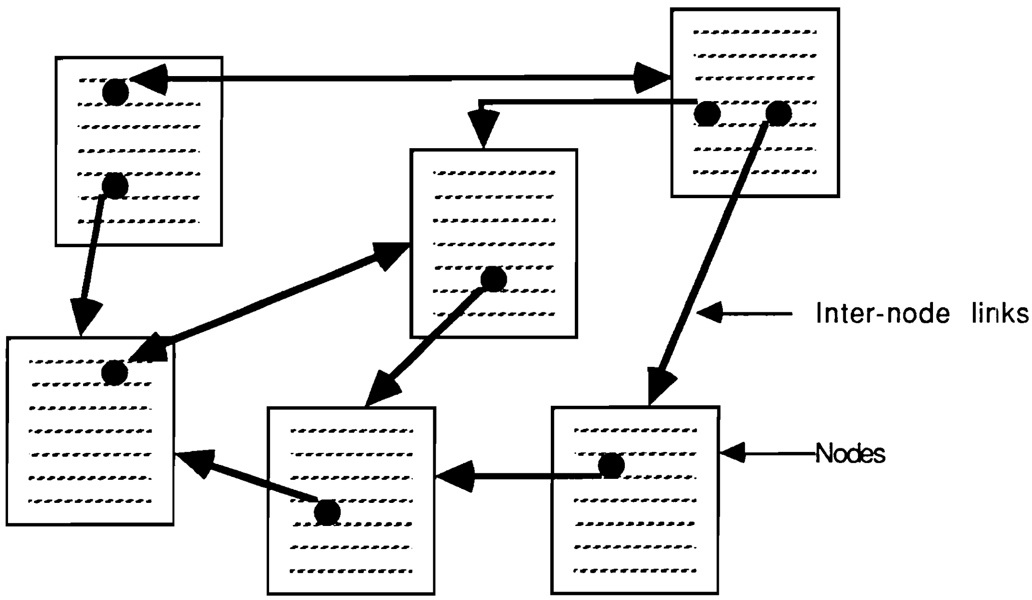


Figure 1. Graphic illustration of a portion of a hypertext system. The system nodes are connected using machine-supported links.

The links may connect related nodes within the same document or between documents. The important characteristics of hypertext systems are (Conklin, 1987; Halasz, 1988; Smith and Weiss, 1987; Tompa, 1989):

- Information is parsed (modularized) into smaller units called "nodes." Depending on display technology limitations, application, medium, and designers' preferences the nodes may be of various data types and sizes.

- The presence of links (or the ability to create links) between nodes provides users with pre-constrained or random access to the network.

Hypermedia systems are identical to hypertext systems with one major distinction: each node in the network may contain animation, graphics, audio (digitized or synthesized), video, executable computer programs, or other forms of information in addition to text. These multi-media nodes are connected to one another using machine-supported links (Figure 2). In terms of appeal, paper-based documents are not comparable to hypermedia documents due to the power and richness of interaction that hypermedia offers. High speed access and large volume memory, high resolution electronic displays, and stereo sound output capability (as in the NeXT computer; NeXT, 1989) combine synergistically to provide a multi-media and highly interactive environment. The important characteristics of hypermedia systems are:

- Information is parsed into smaller units called "nodes."

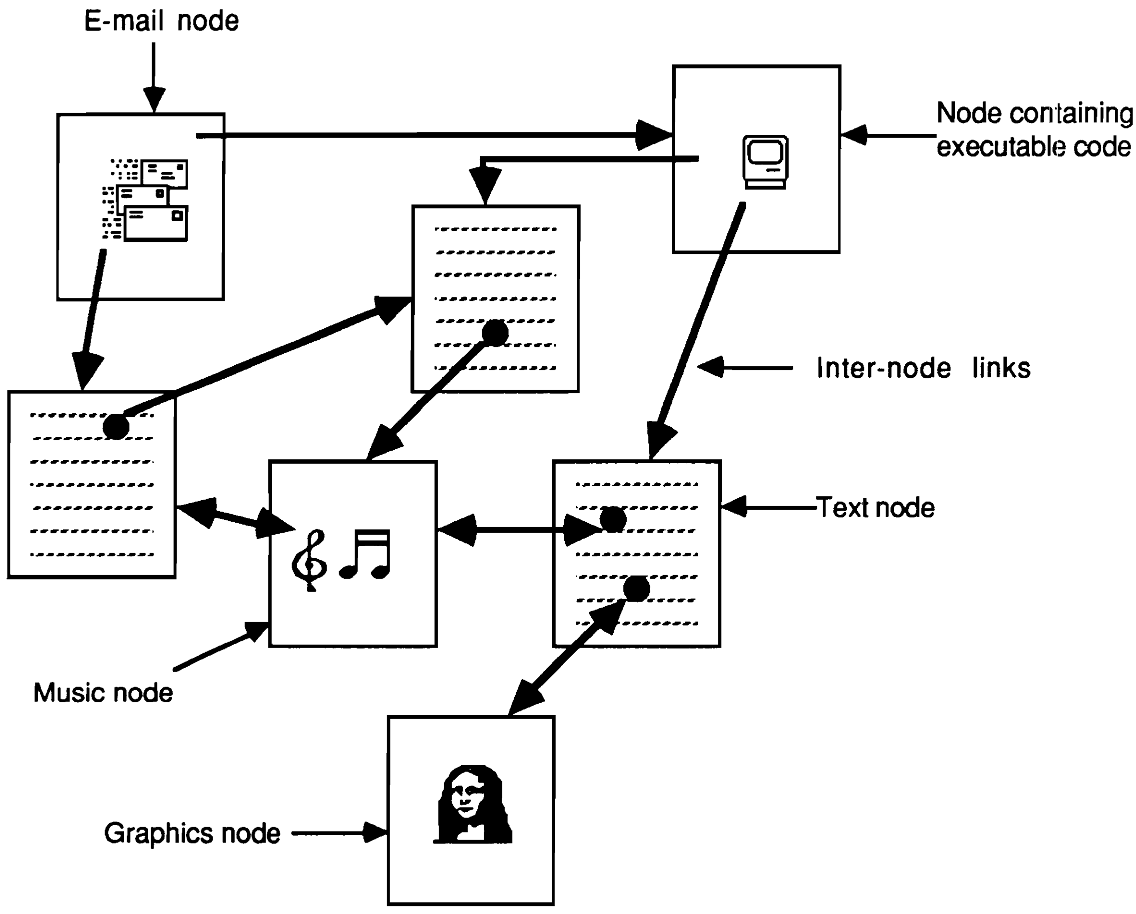


Figure 2. Graphic illustration of a hypermedia system. These systems are similar to hypertext, but include the additional capability of multiple media.

- The presence of links between nodes provides users with pre-constrained or random access to the network.
- There is use of more than one medium (e.g., a system using text, animation, and sound) to convey information.

Examples of hypertext/hypermedia systems include Intermedia (Yankelovich, Haan, Meyrowitz, and Drucker, 1988), NoteCards (Halasz, 1988), KMS (Akscyn, McCracken, and Yoder, 1988), and Apple Computer's HyperCard. The terms hypertext and hypermedia have been used inter-changeably in the literature; however, this research is concerned with only hypertext. While hypermedia is predicted to be the logical successor to hypertext, most of the usability and interface issues for the latter system have yet to be investigated and resolved.

Hypertext Modes of Operation

There are two modes of operation for hypertext systems which should be recognized: (1) information read-only ("browsing") systems, and (2) authoring systems.

Read-only systems. Read-only hypertext systems have been configured exclusively for delivering information. Users may not edit, amend, or in any way alter the information contained in the system. However, this limitation is not applied consistently across all hypertext systems. Often neither the nodes (the information) nor their corresponding links are editable, while more flexible systems allow users to dynamically alter relational links or impose various predetermined sets of links. The former configuration restricts the flexibility and richness of associations between ideas. Users are forced to follow pre-established links, which may or may not correspond to their points of view or perception of related concepts. However, users of these systems are less prone to becoming disoriented in the information space due to restrictions on navigational freedom, and they need not expend cognitive effort in determining which concepts or nodes should be linked. Less flexible systems are appropriate for relatively confined and limited (in scope) applications where the users' needs, informational requirements, and tendencies are well understood or easily predictable.

High-end, flexible, read-only systems (e.g., Intermedia) provide users with the option of tailoring the information to meet their particular needs. A commonly cited benefit of hypertext systems is this ability to configure system structure according to users' interests or requirements, thus creating a richly linked and user-specified information space. Unfortunately, many users are not experienced in database manipulation, and providing mechanisms to create links increases the cognitive overhead of using the system. Additionally, allowing users to browse freely through the information space and travel to any node upon request will, most certainly, lead to severe navigational and orientation difficulties. These systems are considered useful, however, for applications where users are relatively computer literate and are knowledgeable regarding the system and its domain (i.e., the information in the system).

The goal of read-only systems is to provide interested parties with a hypertext environment which simply delivers information to users. The contention is that the delivery of the information will be enhanced by using an extremely interactive, multimedia, flexible system.

Authoring systems. An authoring hypertext system is used for the creation of documents. Users of these systems have the power to create, edit, and otherwise manipulate both nodes (information therein) and links. There are two general classes of tasks performed on these systems: (1) authoring systems may be used to implement an existing linear document in hypertext format, or (2) create or edit a hypertext document (individually or cooperatively with other authors). Many authoring systems can be configured to restrict other users to the read-only mode; however, users of authoring systems usually require the ability to create, edit, and amend documents as they see fit. The goal of these systems is to provide authors with a rich environment in which the structuring of information is highly flexible and where comments, annotations, and criticisms can be made with little or no alteration of the original material.

Computer-Based vs. Paper-Based Documents

Paper-based documents are the most common form of creation and dissemination of information. Hence, it is reasonable to question the utility of moving to an alternate medium, such as computer-based hypertext, for the dissemination of information. Most tasks have been performed relatively well using paper-based media (the "if it ain't broke, don't fix it" attitude). However, the utility and convenience of online information are becoming widely understood and accepted due to the changing nature of information and tasks. While it is important to move forward with modernization and progress, it is unwise to disregard the benefits of traditional methods of task performance.

Since hypertext and hypermedia could be considered the newest trend in online documentation, a review of the inherent characteristics of paper- and computer-based (online) media is appropriate. There are distinct advantages and disadvantages associated with both paper-based and online documents.

Advantages of paper-based documents. Paper-based documents are extremely common and quite popular for many applications. The advantages of such media include:

- Familiarity. Most everyone is familiar with the basic structure and operation of conventional books, newspapers, magazines, and other print media. All users understand how to locate and extract information from these media.
- Higher information density. The amount of material on an average page of text is more than that offered by an average computer screen. Additionally, a quick glance at a paper document provides a good deal of information regarding size and the relative arrangement of sections.
- Portability. Paper-based documents can be transported to and used in most locations. Online documentation requires the use of a computer, which could be an inconvenience if portability is an important consideration for the material.
- Locational cues. Paper documents provide a great deal of feedback simply due to their physical characteristics. For instance, when reading a book, the user can easily approximate the amount of information or length of the document by looking at the thickness and font size (type

size) of the book. Additionally, it is quite easy for a reader to determine what amount of material has been read, where the beginning and end of the document are, or to mark a page of interest (using an "ear-mark" or book marker). Many of these cues are not available in online documents.

- Easy to read. Although technology is rapidly advancing the levels of image quality achievable in electronic visual displays, commonly affordable paper-based documents (e.g., books) have better resolution and image quality than most commonly affordable electronic displays (Dillon, McKnight, and Richardson, 1988; Gould and Grischkowsky, 1984). Unless electronic displays with image quality comparable to paper become less costly, paper-based materials will maintain their lead as the medium of choice for most text-based information (Gould, Alfaro, Barnes, Finn, Grischkowsky, and Minuto 1987; Gould, Alfaro, Finn, Haupt, and Minuto 1987; Jorna, 1989).

- Hardware independence. Once paper-based documents are printed they are accessible and ready for use. Conversely, electronic documents necessarily rely on the availability of appropriate hardware. This factor is also significant for the portability of paper documents.

Disadvantages of paper-based documents. While the advantages of paper-based materials are considerable, there are also significant associated disadvantages. The disadvantages of paper-based documents are:

- Users cannot alter content. Printed documents have fixed content and users are unable to alter the information therein. Paper-based documents must be edited by the publisher (or printer) and reproduced in order to effect any change. This disadvantage is particularly acute for rapidly changing or advancing topics such as current events (history) and science.

- Users cannot customize the organization. Paper documents are static in the sense that the arrangement and order of presentation of material is fixed. Readers are forced to follow the author-defined arrangement, and deviations often result in incomplete understanding of the information (Yankelovich, Meyrowitz, and van Dam, 1985). For many expert users of the document a more flexible and personalized organization may be preferable, which paper cannot offer.

- Costly to produce and reproduce. Labor, paper, printer operating time, ink, and other materials required for the production of paper-based documents are expensive. Moreover, the printing process is a time consuming operation. Thus, paper documents are costly to produce (and reproduce) even when the labor and materials are available. In addition, the current concerns about environmental abuse (destruction of trees) provide more incentive to make use of electronic documents wherever possible.

- Stagnant and quickly outdated. A good portion of information (e.g., science and political events) is constantly altered or updated, and paper documents can become quickly outdated. As previously mentioned, altering the contents of a paper-based document can be difficult and costly; thus, quite often, paper documents are not as current as one would prefer.

- Limited to 2-dimensional representation. Paper documents cannot offer the interactivity of computer systems. For instance, computers can dynamically present the growth of an embryo over time or use 3-dimensional displays for information presentation, while a paper-based document is restricted to static 2-dimensional pictures.

- Restricted to static text and graphics. Paper documents are unable to match the audiovisual capabilities of computers. While paper presents material using static text and graphics, computers offer sound, animation, high speed calculation, and an impressive degree of interactivity.

Advantages of computer-based documents. Online (computer-based) documentation seems to provide solutions to the discussed shortcomings of paper-based documents. Advantages of computer-based materials include:

- Speed of access. Online documents are quickly retrieved by simply accessing a file. As processing speeds of CPUs become faster, the speed of access will increase accordingly.

- Storage and space benefits. Online documents are stored on floppy disks, hard disks, optical disks, tapes, and other compact electronic information storage devices. For example, a 5-1/2-in. optical disk can store 512 megabytes of information (equivalent to thousands of pages of text). Paper documents are stored in specific physical locations and their sheer bulk requires significant storage space.

- Accuracy of information. Unlike paper, online documentation is easily updated, and alterations (both local and document-wide) are easily implemented. As a result, updating information and maintaining its accuracy are facilitated considerably.

- Interactivity and multi-media capability. Computers can be highly interactive devices. This interactivity may be beneficially utilized in information presentation. For instance, a mathematics student can alter the values of variables in a mathematical function and the computer can dynamically alter the corresponding plot. Similarly, a music student reading about the life of Bach may be inclined to listen to one of this Austrian composer's symphonies. Properly equipped computers can play the music and provide for a richer experience. This type of interactivity is impossible in paper-based documents.

- Simple and inexpensive reproduction. Electronic documents are easily reproduced. Duplicating the information in a document is as simple as executing a "copy" command. With computer-based documents, the entire contents of even large scale documents can be duplicated in a matter of seconds, at relatively low cost.

- High degree of connectivity. Today's network linking and communications software provide for an entire network of connected users. Networked users can distribute information through the network at extremely high data transfer rates (compared to paper versions).

Disadvantages of computer-based documents. As with all media, computer-based media have specific shortcomings as well. Disadvantages of online documents include:

- Lack of locational cues. Computer-based systems provide a viewing window (the display unit) from which users must observe the document. Users are likely to lose their sense of orientation in the document since most displays use a 20-line window to move about the document with very few locational cues. Although much discussion and research has focused on this issue (Conklin, 1987; Furnas, 1986; Jones, 1987), no definitive solution has been reached. Potential methods of providing online locational cues include scrolling bars, relative position indicators, and graphical maps.

- Version maintenance. Users of online documentation must have a sense of history and make a conscious effort to maintain copies of previous versions of their documents (or work). Most computers have no "versioning" mechanism, and even when old versions are saved it is often difficult to ascertain where alterations have been implemented (Halasz, 1988).

- Physiological discomfort and fatigue. Complaints of eye-strain, headaches, eye "aches," and general fatigue are common among frequent users of computer systems configured with CRTs. While many standards (e.g., Human Factors Society, 1988) and regulations have been instituted to curb these problems, they continue to be pervasive issues in the work place.

- Difficult to read (poor image quality). As mentioned earlier, paper usually offers better image quality than a commonly affordable electronic visual display. Display technology has closed the gap between these two media; however, paper-based documents continue to be easier to read and significantly less expensive than commonly available electronic displays.

- Hardware dependence. Online documentation requires the use of a computer and the proper accessories. Without these devices, the online document cannot serve its purpose. Eventually, personal computers will become as common as hand-held calculators; however, until then, the hardware dependence of online documentation must be considered a significant disadvantage of this medium.

Task requirements will be the final guide for selecting the appropriate medium. The objective should be to select a medium which best accomodates task requirements, while reducing the maladaptive consequences of associated disadvantages. Nonetheless, it is obvious from reviewing the previous sections that online materials can have important task-related advantages over paper-based media.

Differences Between Current Online Systems and Hypertext

While hypertext is an online medium, distinctions could be drawn between it and a variety of other online information retrieval and delivery systems.

Databases, basic multi-media systems, and multi-window environments share many aspects of hypertext functionality. However, hypertext systems should be differentiated from the traditional systems since hypertext constitutes an extension of these traditional online information management systems, the essential difference being the linking of related concepts (both semantically and organizationally). The differences delineated below are intended to define more clearly hypertext systems through a simple comparison with other environments. The differences are as follows (Nielsen, 1989):

Databases vs. hypertext. In database systems, individual record elements are not self-describing and a specific unalterable information structure is imposed. For example, relational information structures are commonly used in traditional databases. This information structure is necessary to identify records. However, hypertext systems need not necessarily maintain a specific structure; in fact, information in hypertext systems can be unstructured and various users may apply structures which connect information according to their idiosyncratic understanding or need for information. Individual records or information units in these systems can be self-describing and are not dependent on particular relational schemes for identification. Units of information in hypertext (hypermedia) systems are linked solely due to their semantic relationships.

Multi-media vs. hypertext. The essential difference between these two environments is the lack of dynamic movement in simple multimedia systems. Most, if not all, multimedia systems make use of organizational links. These links simply connect information in a specific organization, such as a hierarchical organization. Organizational links perform quite well for multimedia demonstrations (presentations) or as fixed-path educational tools. However, multimedia lacks the power and appeal of hypertext. Hypertext systems allow for rich relational links (which capture the relationships among ideas) in addition to simple organizational links. Ideal hypertext systems also allow users to navigate randomly about the multimedia nodes of the information network whereas navigation in multi-media systems is severely restricted.

Multi-window vs. hypertext. Multi-window systems allow for viewing various parts of a document or of multiple documents on the same screen (also known as space multiplexing). These systems are quite helpful for multi-tasking situations or for expert users who require access to more than one piece of information at a given moment. Often, each window is used to display the contents of a different document and no machine supported connections exist between these documents (or windows). Hypertext systems utilize this type of multi-window functionality, where each node in the network is represented by a separate window (some hypermedia systems do not use multiple windows, e.g. KMS). Simple multi-window functionalities are enhanced in hypertext by incorporating relational links and allowing for dynamic connections between the windows. Multi-window systems require users to access windows individually; however, hypermedia systems provide machine-supported linkages to various pieces of information, which may appear in separate windows.

The above discussion emphasized the differences between hypertext and various other environments; however, there is considerably more similarity than dissimilarity among these systems. Hypertext environments include most aspects of databases, outline processors, multimedia systems, and multi-window systems. For example, most hypertext systems claim to provide information which is independent of organization. However, this type of non-existent information topology may be somewhat impractical in many task situations; designers have resorted to some forms of structured organization similar to relational or hierarchical databases. Many hypertext systems utilize more than one medium and, logically, have much in common with multi-media systems in terms of superficial appearance. Hypertext environments often use many windows to represent various portions of documents; the use of more than a single window is also a major feature of multi-window systems. Clearly, the functionalities of hypertext and these systems dovetail quite well with one another. Hypertext gains its supposed benefits by combining the advantages of these systems and enhances them by providing machine-supported relational links and, in hypothetically ideal systems, random access to any piece of information.

Historical and Logical Background of Hypertext

There are several criticisms and problems associated with current approaches to information management (for both paper-based and online materials). In the opinion of hypertext advocates these problems require solutions, for which hypertext represents the answer. The purpose of the following sections is to present the position of hypertext enthusiasts and elucidate claims regarding the advantageous features of hypertext.

The early vision. The concept of hypermedia is not new. In 1945, Vannevar Bush discussed the problem of document management in the scientific community (Bush, 1945; cited by Bush, 1967). He indicated methods for organizing, retrieving, and reading documents were quite crude, and that it would be impossible to keep pace with the phenomenal expansion rate of human knowledge. Bush suggested a system, called "memex," which would alleviate the problems of both document (information) access and document management. The following excerpt from Bush (1945) illustrates his revolutionary concept:

> Consider a future device for individual use, which is a sort of mechanized private file and library [a memex].... A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.... Most of the memex contents are purchased on tape ready for insertion [a modern day correlate of 'on tape' may be 'on disk' or 'online']. Books of all sorts, pictures, current periodicals, newspapers, are thus obtained and dropped into place [mostly in microfilm format at that time]. And there is a provision for direct entry.... (pg. 27)

Bush continued by suggesting the concepts of efficient storage of information, easy access of material through indexing, and the notion of following "trails" of intellectual interest. These trails are the 1945 version of modern-day hypermedia links.

Bush's main criticisms of the information management resources of the time were their lack of cohesion, lack of connectivity, and the fact that scientific information was basically an "undifferentiated mass." The memex was to be a total system for the creation and management of information, and was proposed to solve those problems. The long term vision was for all writers and information seekers to have access to a machine which would provide them with logical and easy access to the wealth of information in the system. Users could arrange information in any fashion suited to their particular needs. The system was not to discriminate among different forms of data. Sound, text, graphics, and all formats of information were processable.

With the advent of the computer Bush's vision seemed within grasp. In 1967, Bush claimed that the powerful instruments developed in two decades of advancement had made the creation of memex attainable (Bush, 1967). However, the proliferation of computers did not solve the problems of information management; moreover, it seemed as usage and understanding of computers accelerated so did the problems.

Hypertext advantages. The following are most commonly cited benefits of hypertext.

- Connectivity. Theodor Nelson coined the term "hypertext" in 1965. Nelson's concept of hypertext may be described as a computerization of Bush's memex (Nelson, 1967). While Bush intended the memex for a personal workstation and microfilm-based environment, Nelson adapted the idea to computer technology and further advanced the memex notion by suggesting a world-wide literature system. Nelson's criticism of modern information management was the lack of connectivity among various components of literature. Nelson defined a literature to be a system (or body) of interconnected writings, as compared to the traditional "belles lettres" literature. He provided an example of his concept of literature by equating it to the definition implied in the familiar graduate school question, "Have you checked the literature?" Nelson contended all writings are part of a literature and no isolated documents exist; in addition, he believed this inter-connection to be the correct mode for the existence and operation of literature. Thus, hypertext, which by definition is the linkage of material, was perfectly suited for creating, disseminating, and retrieving literature.

- Non-linearity. A major motivation for modern day hypertext/hypermedia systems came from challenging the notion that documents should be presented in a linear format. A linear document is organized as follows: a general overview is presented initially, followed by a more detailed introduction, after which the first section is presented, then the second section, and so on until the final remarks or conclusions are made. Most expository documents are arranged in this fashion such that the logical flow of information moves in a linear sequence from the beginning to the end (this is especially true of paper-based documents). Creators of documents use this linear physical organization to present the logical flow of the document. These documents necessarily couple the logical flow and linear structure of information (Smith and Weiss, 1987). Readers must follow the linear structure (or layout) of the document to understand the material. Deviations from following this linear structure may negatively influence understanding of the information (or logical flow).

Critics of traditional information management denounce the linear organization of documents, claiming it is a limitation of the medium rather than an optimal method for creating and/or presenting documents and information (Bush, 1967; Conklin, 1987; Nelson, 1967; 1980). Hypertext enthusiasts indicate that, in general terms, the logical flow of information need not follow a linear structure. That is, the physical layout of a document should not determine the order of accessed information, but rather users should determine that order based on their interests and needs. These issues are central to the arguments put forth by critics of traditional information management. However, the authors fail to support their statements with empirical evidence. Their contention is that if one considers actual uses of information, it is intuitively obvious that the previously mentioned problems exist (or that the need for this particular functionality exists). For instance, Nelson (1980, pg. 3) writes:

> Consider how it works in science. A genetic theorist, say, reads current writings in the journals. These refer back, explicitly, to other writings; if he chooses to question the sources or review their meanings, he is following *links* as he gets the book and refers to it. He may correspond with colleagues, mentioning in his letters what he has read, and receive replies suggesting that he read other things. (Again, the letters are implicitly connected to these other writings by implicit links.) Seeking to refresh his ideas, he goes back to Darwin, and also derives inspiration from other things he reads -- the Bible, science fiction. These are linked to his work in his mind.

Nelson and his contemporaries support the widespread use of hypertext (and hypermedia) to rectify the problems with existing methods of information management.

Summary

Nelson's (1980) consideration of task performance, while far from a formal task analysis, is intuitively reasonable. At least, Nelson's statements certainly are valid for situations of scientific inquiry. However, the true utility of such systems has yet to be empirically substantiated.

LITERATURE REVIEW

This section delineates a variety of factors and variables influencing the usability, effectiveness, and user satisfaction of hypertext systems. These systems will be generally well accepted by users only if they allow for human-computer interaction (HCI) goals of effective task performance, meet usability requirements, and are satisfying to users (Shneiderman, 1987). A brief description of each of these goals is presented below, followed by a detailed discussion of variables influencing the use (and thus the usability, effectiveness, and user satisfaction) of hypertext systems. Unfortunately, due to the novelty of hypertext research, a majority of these variables lack empirical substantiation and most discussion concerning their influence on usability is purely scientific speculation. Empirical evidence will be presented where appropriate and available.

Human-Computer Interaction Goals for System Usage

Table 1 presents the HCI goals for system usage along with definition and measurements of each. A well designed system must meet certain usage goals or exhibit superior performance (compared to competing designs) along these dimensions.

Ease of learning. Ease of learning in HCI relates the ease and speed with which users reach a specified level of performance (e.g., asymptotic performance). Reduction in task completion times and errors over repeated exposures are frequently used measures of learning. Marked improvement in

TABLE 1. HCI Goals and the Definitions and Measures of Each

CONSTRUCT	QULITATIVE DEFINITION	QUANTITATIVE MEASURE
EASE OF LEARNING	The ease and speed with which users are able to reach acceptable, proficient, or asymptotic performance.	1. Reduction in task time over multiple trials. 2. Reduction in errors over multiple trials. 3. Significant and rapid increase in efficiency or quality of output over time.
USABILITY	"The process of use, the steps the user must take; how the user achieves results." Bennett (1984)	1. Task completion times. 2. Errors. 3. Latency per screen 4. Other measures meaningfully and/or usefully applied to the system of interest.
SATISFACTION	Whether the system generates positive feelings of success, competence, and control.	1. No physical measures 2. Subjective measurement (e.g., questionnaires, rating, ranking, interviews)

quality of work with prolonged system use is an additional measure of ease of learning to use a particular system.

Usability. Usability is probably the most common goal associated with humans' usage of computers. Bennett (1984) has described usability as being concerned "with the process of use, the steps the user must take; how the user achieves results." Gould and Lewis (1985), Roberts and Moran (1983), and Shackel (1984) have indicated the need for quantitative measures of usability. These authors are in general agreement that ease of learning (as measured by the time to learn to use a system to a specified level of proficiency) and throughput (measured by task completion times, errors, and other objective variables) serve as objective measures of usability.

User satisfaction. User satisfaction usually is considered a component of usability, but for our purposes it will be discussed separately. User satisfaction is concerned with whether the system generates positive feelings of success, competence, and control in the user (Shneiderman, 1987). It may be helpful to think of achieving user satisfaction as the minimization of frustration, confusion, and stress. No methods of physically measuring satisfaction exist; however, subjective measurements, such as system ratings, rankings, and verbal protocol data have been used regularly to measure user satisfaction.

Significant Issues In the Creation of Hypertext Documents

Document / data parsing. An essential characteristic of hypertext systems is the parsing of information into nodes. Parsing in this context refers to the modularization of linear documents into separate chunks (or nodes). These distinct units of chunked information are referred to as "nodes" and are relatively modular; thus, parsing of material from existing linear documents is a non-trivial aspect of hypertext system design and usage.

Original creation (authoring) of a hypertext document places the burden of node creation on the author. In most cases, authors make decisions concerning the modularization of information based on their understanding of the domain and user tasks; however, original hypertext creation is not common. Most current hypertext systems are the result of transforming existing linear documents into hypertext format.

Unfortunately, there is no generally accepted method for information parsing. The reasons for such lack of uniform methodology may be due to the novelty of hypertext and the inherent difficulties of applying standardized techniques to relatively amorphous domains. First, the hypertext approach, while far from novel, is still in its stages of infancy. This lack of maturity is exemplified by the fact that commonly available hypertext systems were not marketed until early 1987, and the first professional meeting devoted to these systems finally took place in November of 1987 (Hypertext '87, at the University of North Carolina). Thus, the field is young and many methodological questions have yet to be resolved. Secondly, a standard approach to information parsing may be impractical due to the wide variety of topics and areas of information which could potentially be converted to hypertext. It will be difficult to develop such methods without proper understanding of the information to be parsed. Nonetheless, many have attempted to approach the problem of parsing from a variety of viewpoints. Generally, there are two types of data parsing approaches: quantitative and qualitative.

In its most crude form, quantitative data parsing is performed by selecting a reasonable number of words (or characters), *x*, to constitute a single node. Thus, node 1 would contain words 1 through *x*, node 2 would contain words *x* +1 through 2*x*, and so on. This method can be automated and simply executed given a proper parsing grammar; however, it does not seem to be an elegant technique for data parsing and may not provide for beneficial utilization of nodes.

Others have used qualitative approaches. For instance, the original document could serve as the guide to the parsing process. Every chapter, section, and subsection is considered to be a separate node. This method imposes a somewhat hierarchical structure (e.g., chapter to section to subsection) and, accordingly, links are created to allow for navigation through this strict hierarchy of nodes. A second qualitative approach is the use of subject matter experts (SMEs). SMEs are an excellent source of assistance since they are familiar with both the information domain and its applications. SMEs can provide guidance in determining the information to be contained in nodes, node size, and node linking. Document parsing and node linking are closely related, and under ideal conditions each task should be performed with regard for the other. Most current systems rely on qualitative parsing of documents, mainly due to the lack of other methods.

Links. A link establishes a connection between two nodes in a hypertext network. Links provide paths by which users can retrieve or access information in the network (once the material is parsed). The essential component of linking, which differentiates hypertext from other forms of information management, is that links provide for non-linear access to material. Hence, links are more than simple page turners used to move from the current position to the next (as in a linear online document); links allow users to actively select the topic or portion of information to be viewed or accessed. Hypertext advocates (Conklin, 1987; Halasz, 1988) indicate that linking is the essential component in hypertext systems.

Links have been used to connect related information; these relationships between nodes of information can be of many sorts. That is, as long as some useful relationship exists between two nodes they may be connected. In general, two distinct but related types of linking have been identified by Conklin (1987) and accepted by most hypertext professionals (Halasz, 1988): organizational links and relational links. Some hypertext systems use both linking types, while others use only one of the two. The following sections describe and classify the link styles and attributes.

- Organizational links. Organizational links provide a mechanism for the arrangement of nodes through the use of traditional database organizational schemes. These schemes include, but are not limited to, hierarchical, alphabetical, and functional organizations. By far, the most popular method for such organization is the hierarchical scheme. Hierarchies are designed such that users enter a document or system at the highest level of generality and proceed to consecutively more specific nodes of information. For instance, a geography database might attempt to maintain spatial relationships existing in the physical world. The user would enter the database at the most generic concept: the planet earth. Consecutively higher levels of specificity are accessed by retrieving information on continents, countries within the continents, provinces or states within countries, cities within provinces, and so on. Conversely, the same database could employ a different hierarchical relationship for information organization. Assume the system is intended for retrieving information on landlocked vs. non-landlocked countries. The strict spatial relationship discussed earlier may not be as useful. Here, the world's countries would be divided into two large groups: landlocked and non-landlocked. Perhaps under the landlocked category, countries would be organized by their distance form the nearest body of water whereas, in the non-landlocked category, countries may be organized by the number of different bodies of water they border. Hence, there is a variety of dimensions one could use in creating hierarchies. Ideally such hierarchical organizations should be based on the users' needs for and uses of the information.

Classical hierarchical configurations allow for movement in only two directions: to higher levels of specificity or lower levels of specificity. Figure 3 illustrates a pure hierarchical design. At each node, users decide to traverse one of the possible links that leads to a node which contains information of a higher level of granularity than the previous node. Hierarchies are limited in flexibility and quite predictable. For instance, assume the top node of a hierarchy to be level 0. If two nodes at levels i and j have a nearest ancestor (or common upper bound) at

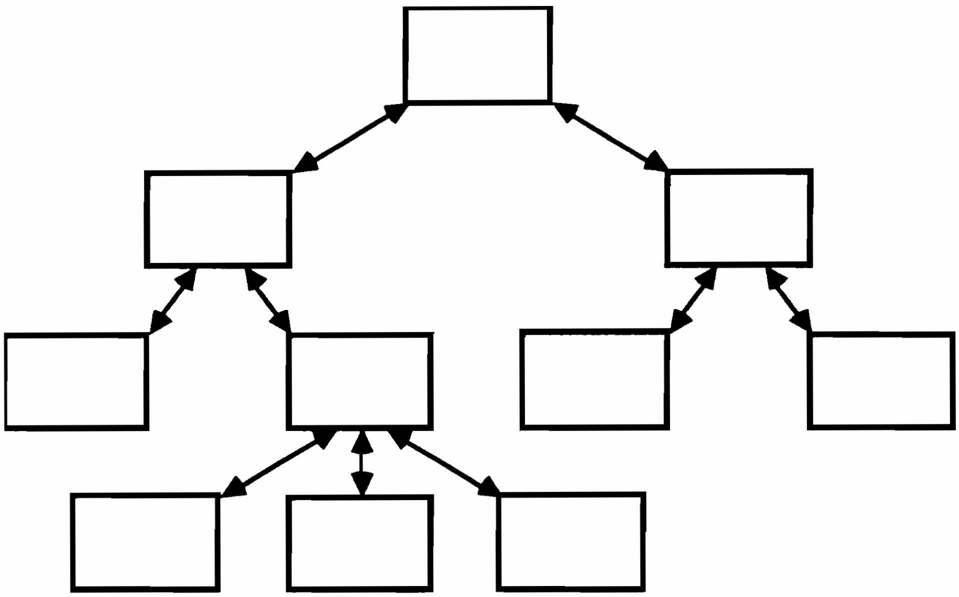


Figure 3. Graphic illustration of a pure hierarchy. Nodes are arranged such that users can access either more specific information, by moving down the hierarchy, or more general information, by moving up the hierarchy.

level k, then the following formula easily predicts the distance between the nodes:

$$N = (i - k) + (j - k) \quad (1)$$

where N is the number of traversed links between nodes i and j (Parunak, 1989). Examples of organizational links include:

1. links from the current section to a subsection (or vice versa), and
2. links connecting two (or more) successive nodes.

Some authors contend such hierarchies with limited flexibility are no different from a menu-driven hierarchical system and do not embody the true essence of hypermedia. It is not the intention of this study to resolve the issue of whether purely hierarchical systems should be taxonomically different from other hypermedia systems; rather, it is sufficient to realize that hierarchical organizational schemes are quite common for organizing networks, regardless of the class of information systems to which they may belong.

- Relational links. Relational links are established on the basis of semantic relationships among the nodes of information. In this context a semantic relationship is simply any meaningful explicit or implicit relationship between two nodes of information. It is helpful to consider relational links in terms of references from one node to another: the origination of the link is the link source and is the reference to a word or region (or entire node) in another node, which is the referent. A relational link simply establishes a connection between the reference, in one node, and the referent, in another node, with no regard for their relative positions in the network (in organizational terms). Figure 4 illustrates such relational links. The "meaningfulness" of the relationships will depend on the information domain and users' task needs for the information, as will the decisions as to which nodes should be linked. There are five dimensions which characterize relational links: (1) bases for establishing links, (2) internal meaning, (3) user- and author-established linking, (4) inter- and intra-document links, and (5) one-to-one vs. multiplexed (one-to-many) linking (Figure 5).

1. Bases for establishing links. An author or user may connect two nodes based on some dimension of interest or utility. Hence, the term "relational link" merely indicates that two chunks of information are linked based on some implicit or explicit relationship, it does not include specific rules for the creation or classification of such links. Examples of bases or dimensions for establishing links include, but are not limited to: supplementary material, explanatory information, definitions, examples, opposing information, comments regarding the current node (information), and links from text to cited graphs or tables. In short, the bases for establishing links are varied. Pragmatically, however, both hypertext system authors and users rely on task- and domain-relevant dimensions for the creation of relational links. Some systems restrict the user to a specific set of link types. For instance, users may be allowed to create links that connect two nodes only if the information in one supports or refutes information in another. The increased organization derived

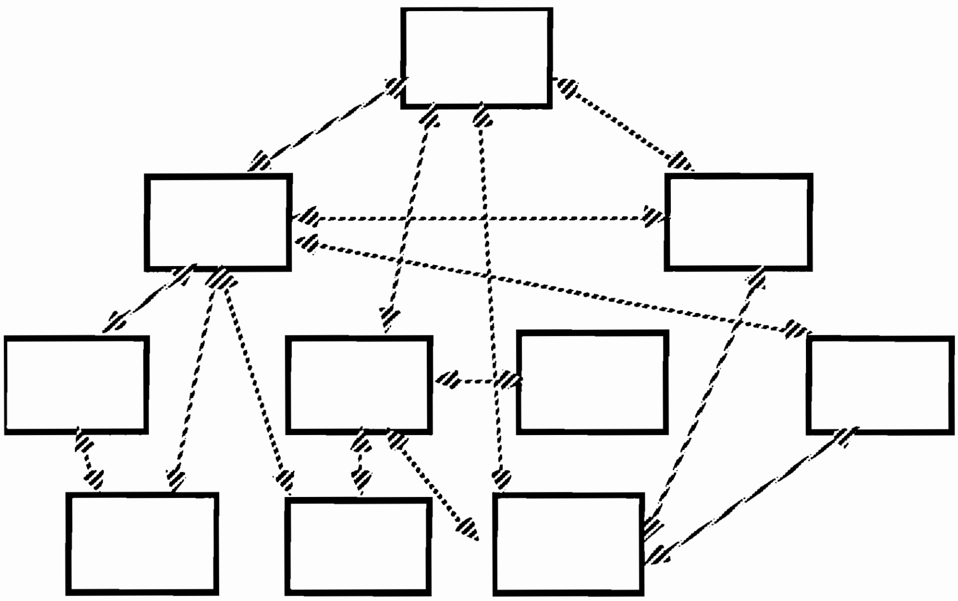


Figure 4. Relational links provide connections among nodes based on some relevant (or task-appropriate) dimension. As seen in the illustration, no consideration is given to the organizational layout of the network; any two nodes which are deemed related are connected.

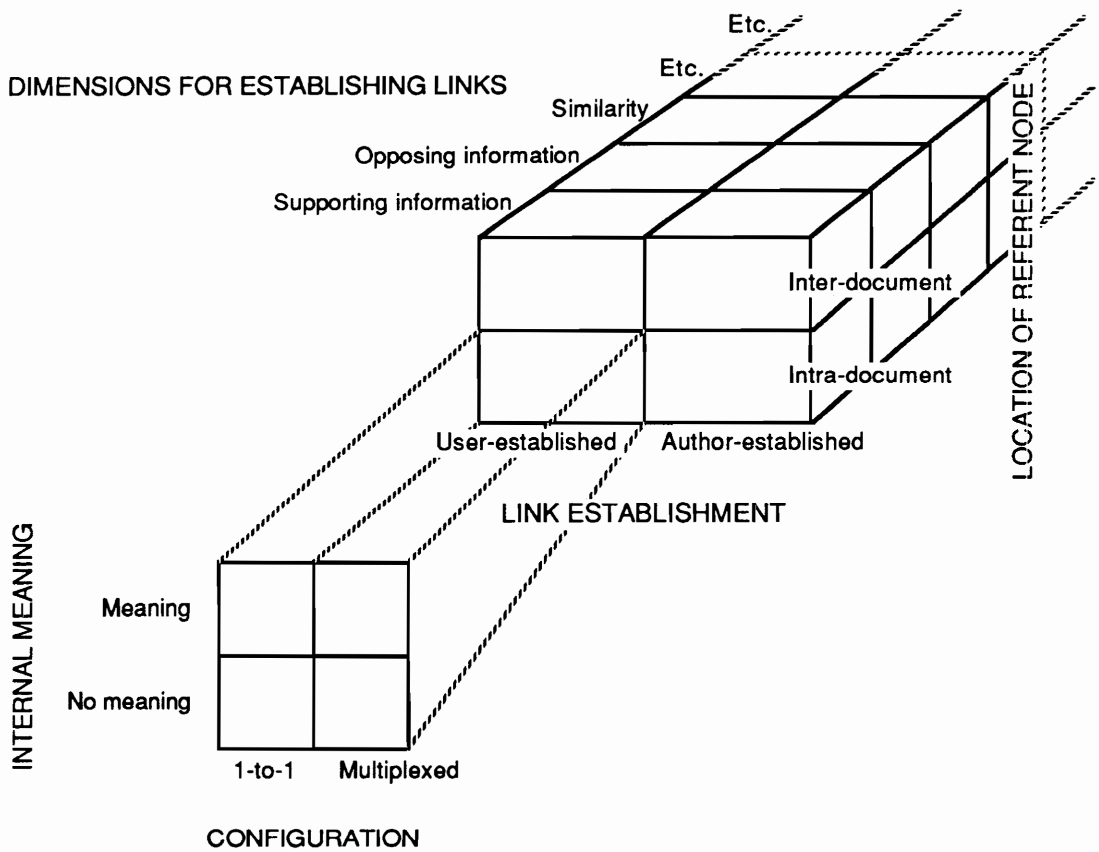


Figure 5. Relevant dimensions by which relational links can be described. These five dimensions may not constitute an exhaustive list of linking factors, but are the most salient.

from such linking restrictions is believed to be an advantage.

2. No internal meaning vs. internal meaning. Relational links are always established between nodes based on the existence of some useful relationship. However, some systems use links which do not provide indications regarding the exact nature of the relationship, while others include links that inform the user as to the type of relationship. In the first case, the user is aware of the link and that its traversal (or activation) will access information related to the current node, but the exact relationship is not apparent until the user actually activates the node. Once at the referent node, the user is assumed to realize the relationship and will decide to continue or to return to the original node. In the latter case, the nature of the relationship is obvious from the link labeling or other provided cues (at the reference node). For example, a link indication may be labeled "Click here for opposing view-point." Users of such a system are aware that clicking on the link indicator will access material which opposes information at the current node. Other such links provide a small window of information regarding the nature of the relationship they represent and/or information concerning the referent node. The user then decides to pursue the link or remain at the current node.

3. Author- vs. user-established links. Relational links may be either author- or user-established. Most currently available information retrieval systems offer only author-established links, while more flexible systems allow for user-generated links. System authors establish relational links among nodes based on their understanding of user needs, task

requirements, and the information domain. Creation of sophisticated systems often requires domain experts to assist system authors in the establishment of such links. Both quantitative and qualitative approaches to link generation have been offered. These methods are discussed later in detail. Many systems with author-established links provide no mechanism for user configuration or editing, yet others provide the flexibility of user-established links. In such systems, users may create links between (among) nodes they consider to be related in some meaningful way. Some hypertext advocates claim all hypertext systems should provide for user-established links, but there are limitations which hinder the common availability of such links. First, many users are not highly familiar with the information domain and, thus, unaware of the potential relationships which could exist. Secondly, the overhead (both in terms of system design and users' cognitive demands) associated with such linking could be unacceptable. Users would be required to learn the commands and operations for establishing links and would need to be aware of their current location and that of the potential referent node in the network. This localization may not be easy if users are navigating about the document with complete freedom; disorientation and frustration may become prevalent. However, user-configurability may allow for more appropriate use of the system since users will be able to establish links which meet their needs and requirements.

Currently available systems provide little user configurability of relational links. Most relational links are imposed by the system authors/designers, who are presumed to provide links which the user can beneficially utilize.

For small applications with relatively few uses the author-established links may be sufficient; however, no evidence exists as to the utility of such links in situations where timely and accurate retrieval of information is necessary. Moreover, unless each user can maintain individualized sets of links (as in *Intermedia*, Yankelovich et al., 1988), user-established links may lead to disorganization and confusion. For instance, links established by one user may seem completely non-intuitive or useless to another.

4. Intra- vs. inter-document (database) links. The preceding discussions have focused on creating links to access different nodes within a document or database and, by far, this type of linking is most common. However, a particular node may be related to information in a separate document, as well as the document to which it belongs. Hence, inter-document relational links may be established in addition to intra-document links, assuming another online database (or document) exists. Most available systems generally cover a single topic or area, but there are systems which contain more than one database. A wider range of topics may provide for a richer information environment. A simple example of this type of linking may be a document which allows users to access referenced documents by clicking on the author's name. However, such complex linking may create problems of navigation and disorientation (Jones, 1987; Nielsen, 1990). Additionally, the number of available links per node may rapidly increase with the inclusion of additional databases. While the optimum number of links per node is not

known (and may be database dependent), it is clear that users should not be inundated with numerous links.

5. One-to-one vs. demultiplexed linking. Relational links may be designed as one-to-one or demultiplexed (one-to-many) structures. A one-to-one link connects only two nodes: from one reference to one referent. However, a demultiplexed link connects a single node to many referent nodes. A one-to-one link is exemplified by a link connecting two nodes because information in one node supports information in the other (Figure 6); selecting the link will directly access the referent node. A demultiplexed link would connect the current nodes to numerous other nodes with supporting information. Once the user activates the link, a selection screen (or menu node) is provided from which the user determines which of the linked nodes is to be accessed (Figure 7). Thus, through the activation of a single link the user has access to different locations throughout the document.

Relational linking is presumed to provide the power of hypermedia (Conklin, 1987), this power being derived from the added flexibility and supposed utility of this type of linking. Additionally, some proponents claim relational links provide an added facet of power because they correspond to human information organization (Fiderio, 1988). Much research, generally based in psychology, has given support to the notion that words are stored and organized in the brain according to their meanings and semantic relationships. Early experimental work in word free-association led Kirkpartick (1894; cited by Solso, 1988) to claim that impressions (e.g., words, pictures, etc.) are stored, recalled, and

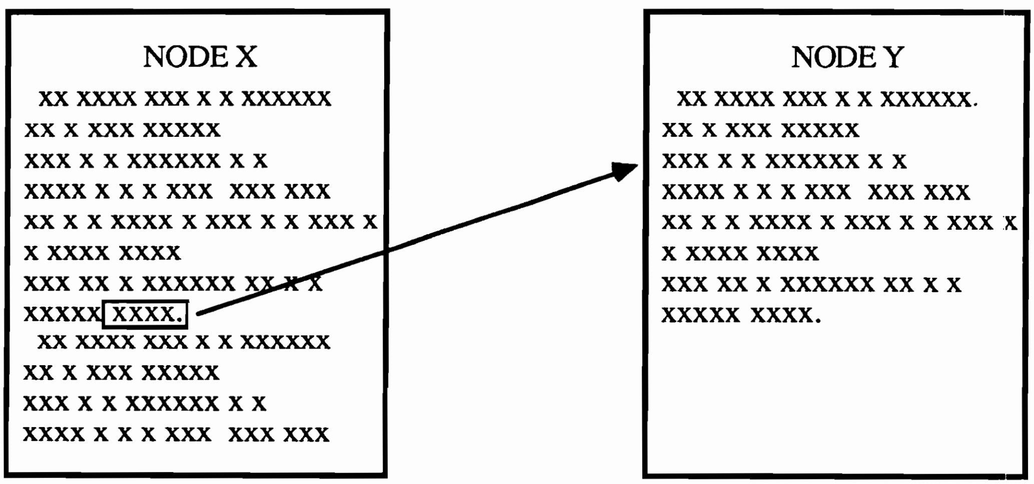


Figure 6. One-to-one relational links.

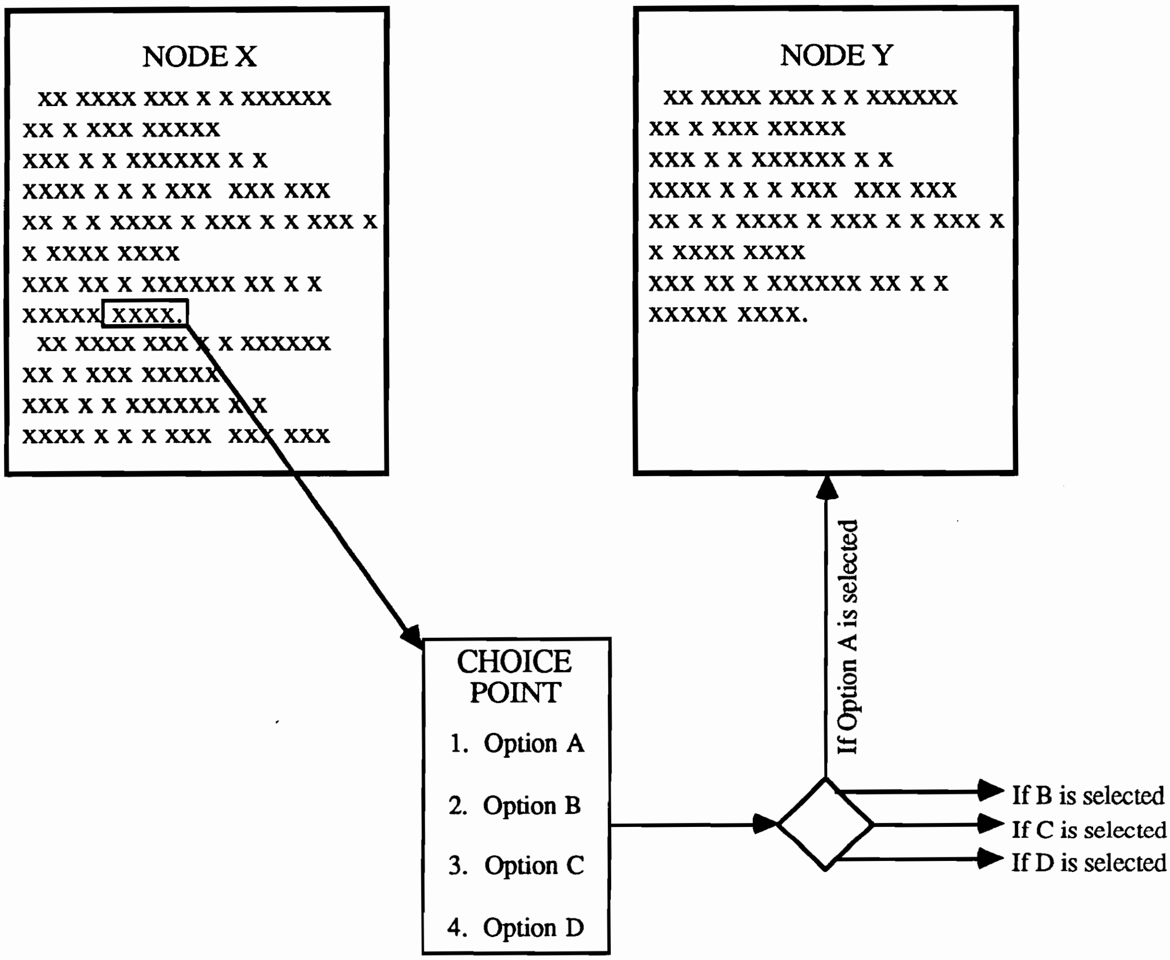


Figure 7. One-to-many (demultiplexed) relational linking.

recognized due to their belonging with other impressions; this research indicated the existence of some form of semantic (meaning-dependent) organization of material. Recent research in cognitive psychology (Bower, 1970; Smith, Shoben, and Rips, 1974) has reconfirmed this notion. This basic work in psychology has led to semantic information organization in a variety of settings, including hypertext.

Linking techniques (methods). Most hypertext systems, especially information retrieval systems, are designed with minimal flexibility and users depend on author-established links. As a result, the focus here will be on author-established links and the techniques for establishing such links, in particular author-established relational links. Hypertext advocates indicate relational links are the most important and useful type of link and that these links, when properly designed, allow for optimal task performance.

In practice, however, linking of information can be somewhat difficult and, at times, nebulous. This section provides additional detailed examples of methods used for determining author-established links. There are both quantitative and qualitative techniques for establishing links. The qualitative techniques seem to be most commonly used since no quantitative method has been empirically validated in more than a single setting, and the generalizability of these methods to multiple information domains is largely unsubstantiated. Thus, usually qualitative methods are used to meet the needs of specific information domains.

- Quantitatively derived links. Furnas (1986) presented a method of connecting information called the "fisheye view." Furnas draws an analogy between a database user's viewing strategy and that of using a fisheye lens. A fisheye lens illustrates physically proximal places in great detail while still presenting information regarding more distant places. However, as the locations increase in distance from the current position, pertinent information is included with progressively less detail. For instance, one standing in New York City would have a wealth of detailed information about New York City, Northern New Jersey, and most of Connecticut. Albany would be represented in lesser detail, yet in more detail than Washington D.C. Denver and San Francisco would be represented in as little detail as possible. The underlying concept of this view is that physically proximal locations are of more interest than distant locations. Furnas claims the notion of degree of interest (DOI) can be employed in information databases. He proposes a DOI function which assigns an "interest" number to each node based on its distance from other nodes of interest (presuming a pre-established organizational structure of some sort) and pre-established importance values for each node. The degree of interest function is:

$$DOI_{fisheye} (x \mid .= y) = API (x) - D (x,y), \quad (2)$$

where $DOI_{fisheye} (x \mid .= y)$ is the user's degree on interest in point *x*, given that the current position is *y*, API (x) is the global a priori importance of *x*, and D (x,y) is the distance between *x* and the current position, *y*. The DOI number indicates the data element's level of interest to the user, given the current position. Hence, the *n* most proximal (in terms of DOI, not physical distance) data elements may be presented as linked items for the current user position.

Lacy, Chignell, and Kinnell (1988) discuss the quantitative establishment of links using statistical criteria. These authors describe building a hypertext document management system. An important consideration was the issue of connecting documents in a meaningful fashion. Lacy et al. used co-occurrence frequencies to determine items to be linked. Their premise was that highly inter-related articles had a higher probability of using similar terminology. Since articles in the database had assigned index terms (one or many), the linking technique can sample these index terms for co-occurrence of words. Criteria can be established such that a particular number of co-occurrences of index terms in articles automatically establishes a link between the two words everywhere in the system. They provide the following example. If the terms "equal opportunity" and "affirmative action" tended to appear together frequently as index terms for articles, then a link was established between them.

Certainly, these methods offer a degree of automation for linking and can be helpful in many situations. However, there are issues which limit their applicability. For instance, in the "fish-eye view" scheme, the important condition for the application of the DOI formulation is that some underlying data structure exists to define the necessary components. Furnas (1986) uses a hierarchical organization to discuss the application of the DOI function. This notion assumes the current organizational scheme is ideal for the data. Secondly, the distance between two nodes (as defined by the system organization or structure) will heavily influence the DOI number achieved between them (between *x* and *y*). Thus, a distant data element must have a very high degree of a priori importance to achieve a reasonable DOI number. In the co-occurrence scheme, problems could occur where co-occurrences of words or terms do not suggest appropriate semantic relationships. The co-occurrence scheme may simply be insensitive to many task-related relationships which do not meet the inclusion criteria.

- Qualitatively derived links. A major proportion of current systems rely on qualitatively derived links. Reliance on qualitative linking is largely due to the absence of a superior alternative. No truly useful quantitative approaches exist, and the general applicability of any quantitative method of linking would be somewhat questionable since information domains vary considerably. Specifically outlined rules for the creation of links do not exist; hence, links are established by system designers based on their understanding of the information and user tasks as determined by task analyses and interviews with subject matter experts. Thus, the supposed uses and needs for the information determine the nodes to be connected. For instance, Shneiderman (1988) describes the HyperTies system used in an information delivery capacity for the University of Maryland's student union. In this system, authors have provided links to various elements of information based on their understanding of the potential needs of users (Figure 8).

INTRODUCTION: STAMP UNION page 1 of 3

The Adele H. Stamp Union, formerly known as the Student Union, is the cultural center for the University. The Union provides a variety of services to the faculty, staff, and students. A plethora of **restaurants** is available providing a wide choice of atmosphere and a variety of menus. The Union is also a center for **entertainment**. Several **shops** and many **special services** are available, too. **Union programs** include concert exhibitions and craft classes. The Union is open all week from 7 AM to 1 AM, Monday through Friday, and until 2 AM on weekends.

NEXT PAGE **END**

Figure 8. Sample screen from the HyperTies Student Union system. The bold type indicates linked words. Activating a link will access information related to the selected word. These links were established based on the designers' understanding of user interests in using the system. For instance, a good deal of information concerning shops in the student union is available; however, a patron interested in reading more about Adele H. Stamp cannot access such information from this node (if at all available). The designers assumed there would be no (or very little) use in including such material since very few people, if any, would require such information from the current node. (Adapted from Shneiderman, 1987.)

Another qualitative scheme is to connect nodes based on designated keywords. Any time reference is made to information contained in another node a link is established. System designers or SMEs may determine the keywords signifying the origination of a link. For instance, an online automobile repair manual may have a node describing the pistons. Links to the "piston" node would be established from any node referring to the pistons, regardless of location or level of detail required.

In practice, most hypertext systems offer both organizational and relational links for the network of nodes (as illustrated in Figure 9). For instance, Conklin (1987) presents linking data from a case study of NoteCards™. NoteCards is a highly powerful, AI workstation-based, hypertext, document creation environment developed by Xerox Corp.. Of the total 3460 links created by the user observed during the case study, roughly 73% were considered organizational (i.e., hierarchical) while 8% were relational, and the remainder were E-mail links. However, these data are from only one user and perhaps unrepresentative, as NoteCards is mostly used for document creation rather than information retrieval. Additionally, users of this particular system are assumed to be computer literate; in fact, the system is targeted toward scientific environments. A similar investigation using an information retrieval hypertext system may yield dissimilar linking tendencies.

Task-Related Variables Influencing Usage of Hypertext Systems

The following task-related variables are postulated to have an effect on hypertext usability.

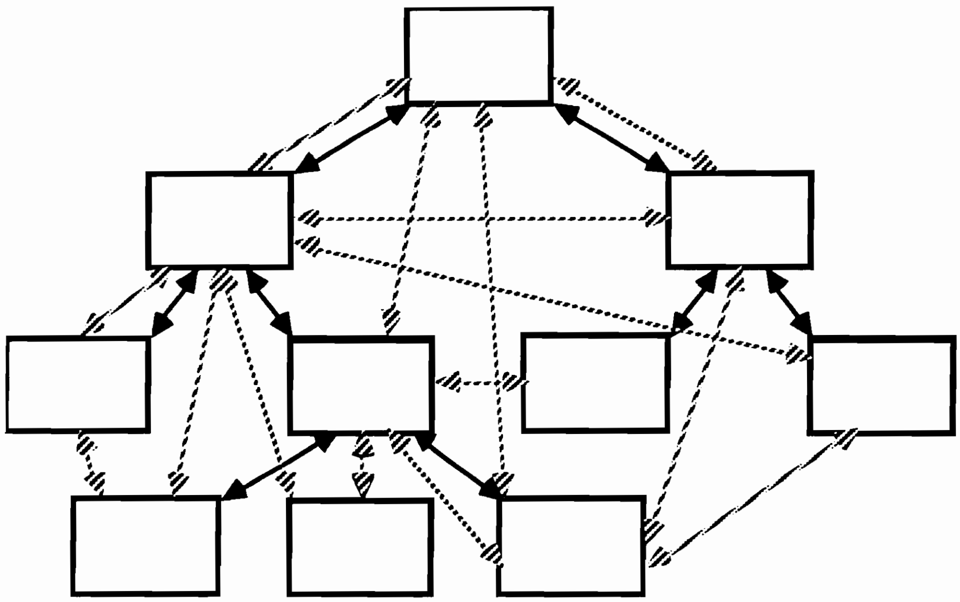


Figure 9. A linking configuration combining both hierarchical and relational links. Many current hypertext systems offer such a network configuration. In the illustration, the solid connections represent hierarchical links while dashed connections represent relational links.

Nature of task/application. Task application refers to the type of task(s) to be performed using the system. There are four general task application areas proposed for hypertext systems (Conklin, 1987; Frisse, 1988; Halasz, 1988; Shneiderman, 1988): (1) information retrieval, which is simply accessing information for one-time use or reference (includes repair/maintenance manuals); (2) learning, which is accessing information for educational or training purposes where retention and or transfer to an applied setting is required; (3) pleasure, which includes entertainment and leisure activities or usages; and (4) document creation and organization, which is actual creation and implementation of hyper-documents (such activities may include programming and design decisions). The tasks performed in each of these areas are quite different in terms of task requirements, error tolerance, informational requirements, and perhaps education and background of system users. This study will concentrate on the information retrieval domain. It is believed that this area will constitute the majority of hypertext applications in the foreseeable future (Brown, 1988; Muto, 1990; Shneiderman, 1988). In fact, the variables to be discussed in the remainder of this chapter will be discussed within the context of information retrieval applications.

Integration of information. While there is a variety of task requirements, perhaps one of the most important is whether tasks require the integration of material read from various locations in a document in order to answer specific questions. For instance, an information retrieval scenario requiring no information integration is exemplified by a case in which the required information is located in one, and only one, node. A task requiring integration is exemplified by a case where the correct answer to a given query (or retrieval problem) is obtained only through combining information from two (or more) nodes. A parallel in a learning application might be where a student must integrate information learned on a variety of nodes.

Commonality between query and language used in the system. The degree of agreement between the language (wording) of a query and that of the information in the system influences the efficacy of system usage. A concrete example of such a condition is where the user of a system must answer a question that may or may not correspond to the organization or wording used by the system author(s). Dumais (1989) and Dumais, Furnas, Landauer, Deerwester, and Harshman (1988) indicate that the vocabulary used in retrieving appropriate documents in a document retrieval system have a profound effect on the number of relevant articles the query yields. Egan, Remde, Landauer, Lochbaum, and Gomez (1989) and Campagnoni and Ehrlich (1989) have investigated the influence of this issue on performance of hypertext systems. Egan et al. compared a document implemented in a hypertext system (designed at Bellcore and called "SuperBook") and a conventional paper-based book. Their comparison tasks involved subjects retrieving answers to questions whose wording was taken from the text and section headings as they appeared in the document, from the text alone, from the headings alone, or taken from neither. Both documents performed better under those conditions where there was some degree of agreement between the question and the information. The hypertext system performed significantly better than the paper-based document for questions which were worded with either text and headings or text alone. In the hypertext system, users accessed a "Word Lookup" (i.e., string search) function. Hence, the system was faster and more accurate than its paper-based counterpart. However, these results may not point to the advantages of hypertext per se, but rather those of computer-based enhancement. Undoubtedly the electronic chip is faster than the human eye in locating a specific string. Nonetheless, the importance of vocabulary commonality is demonstrated.

Time constraints. Time constraints refer to the amount of time allocated for task performance. Marchionini and Shneiderman (1988) suggest that time constraints could have a profound influence on system usability. This hypothesis may be reasonable since time stresses seem to influence user performance with other computer systems. For instance, military, process control, and telephone operator applications may require time stressed task performance, whereas entertainment applications may have no specific time constraints.

Physical environment and location. This factor refers to the location of system use. In most instances, the application is specific to particular environments. Military applications, for instance, are used in a different setting than office system applications. Environmental elements such as glare, ambient illumination, noise, and vibration could conceivably influence system usage.

Information Domain Variables Influencing Hypertext Systems

Clarity of inter- and intra-document relationships. The clarity or explicitness of inter-relationships varies considerably with the information domain (or document) to be implemented in hypertext format. Since hypertext relies on linking and links are meant to capture semantic relationships, it is clear that the explicitness of such relationships will impact usability and design. This is perhaps an area of study for such disciplines as linguistics or English.

Data format of information. The pre-online format of information is important in representing the original document in computer-based systems. Data format refers to the original format of the information, such as photographs, text, sound, etc. For instance, the online presentation of material existing in textual format is not a major technological effort, whereas an acceptable online reproduction of a good quality photograph may require a high resolution, full color display and complex image processing algorithms. Hence, the original format of information and the ability of the hypertext system to reproduce such information accurately and precisely will influence both usage of the system and many design decisions.

Pre-existing information topology. Certain domains may have existing organizational schemes. For example, most phone books are organized alphabetically. This pre-existing topology may have an influence on the usage and design of hypermedia systems used in such domains. These existing topologies may be quite useful in implementing organizational links.

System Design and Configuration Variables

System features. This factor refers to features that distinguish hypermedia systems from one another. As previously mentioned, there are no specific defining characteristics of hypertext systems. Hence, designers of such systems tend to augment basic hypertext concepts (i.e., non-linearity, linking, and modularity of information) with a host of features. Schwalm and Samet (1989) refer to this trend as "Feature Shock." Such features include, but are not limited to, various methods of upgrading information, methods of providing locational cues (e.g., graphical browsers), use of multi-window capabilities, various types of linking mechanisms, embedded and explicit menus, various types of search mechanism provided (e.g., browsing, string searches, Boolean logic search, keyword in context search, indexed word retrieval), and other options. The utility of including such features is unknown, and it is best to include only those features which perform a useful function toward task completion.

Flexibility. Flexibility is the extent to which users may alter the information and links in the system. Read-only hypertext systems vary considerably in the level of allowable flexibility. For example, the HyperTies museum installations at the B'nai B'rith Klutznick Museum in Washington D.C. and the International Center of Photography in New York City (Shneiderman, Brethauer, Plaisant, and Potter, 1988) simply provide users with pertinent historical and background information. Users are presumed to have no need for or desire to alter the system material. Conversely, more flexible systems, such as the Intermedia educational hypermedia system (Yankelovich et al., 1988), allow users to establish personalized links among system nodes; however, no alteration of the information therein is possible. Finally, most authoring systems allow for high levels of flexibility for both linking and information editing.

For most systems, flexibility is determined by the nature and requirements of tasks to be performed. However, it is reasonable to postulate that the level of system flexibility could have an influence on usability and user satisfaction. This effect may be especially pronounced in systems which do not provide the appropriate flexibility for the tasks to be performed. Thus, poor performance and user dissatisfaction are predicted for conditions where the system lacks the flexibility needed to optimize task performance or where unnecessary flexibility is included, resulting in user confusion and frustration in addition to degraded performance.

System size. System size refers to the literal size of the hypertext system as determined by the number of nodes, the number of links, the number of different databases (or topical areas) in the system, and/or any combination of these factors. Cecala, Watkins, and Mohageg (1990), Raymond and Tompa (1987), and Jones (1987) have indicated that system size could influence usability. Perhaps as system size grows or the more databases it includes the more viable and beneficial hypertext approaches become (Figure 10). Unfortunately, there are no empirical data to substantiate these hypotheses.

Distance from target node. The number of nodes through which users must pass to reach their target node(s) may have significant impact on system usage. If users are only 1 link away from the target node no serious complications should arise, unless the user mistakenly accesses a completely irrelevant node and becomes disoriented. However, traversing larger numbers of links may have concomitant negative effects. Navigating through 4, 5, 6, or

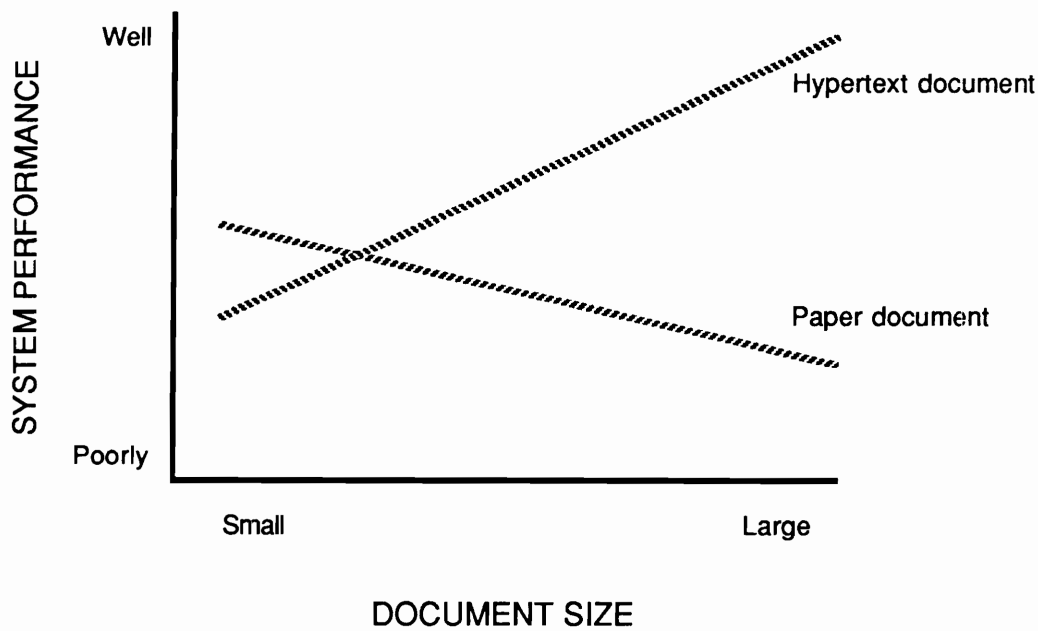


Figure 10. Hypothetical influence of size on system performance. Conceivably, paper-based documents may perform quite well, and indeed better than hypertext, for small to average size documents. However, as document size and complexity increase, perhaps hypertext documents will outperform paper. Unfortunately, there is no empirical support for this hypothesis; in fact, the point at which hypertext begins to surpass paper is not known and may vary among applications.

more links (and nodes) may prove difficult for many users. A user's distance from the target node may be determined by factors such as:

- availability of complimentary search mechanisms,
- user's ability to properly utilize search mechanism,
- efficacy of complimentary search mechanism(s),
- user's position in document at the initiation of the retrieval task, and
- system size (depth/breadth).

Unfortunately, no data are available concerning the relative merits of different systems with respect to the traversal of numerous links in hyper-"environments."

Available data formats in the hypertext (hypermedia) system. This factor refers to the number of different media available for information presentation in a hypertext or hypermedia environment. An often cited advantage of hypertext and hypermedia is the inclusion of multiple media (e.g., text, graphics, sound, animation). Undeniably, multimedia capability is impressive, compelling, and could be helpful in task performance. For instance, using animation to illustrate the changes in a boiler temperature over time may be useful in process control applications. However, optimum media for presentation of information are not always obvious. For instance, information from the boiler temperature example could also be presented in a tabular format. As usual, task requirements may determine presentation medium. If general temperature trends are of interest then graphical representation is appropriate, whereas if exact values are of interest then tabular formats may be preferred. Hence, design decisions regarding the medium of information presentation can be critical. Selection of appropriate media and proper presentation mode for each media will be key if the system incorporates multimedia capabilities.

Linking factors/issues. These factors determine the usability and appropriateness of links (these are not quantitative). Linking is considered to be an integral part of hypertext and, accordingly, issues related to linking influence the usability of such systems. The important aspects of linking will be briefly mentioned here since they were discussed in detail in previous sections.

• Linking structure. Organizational and relational linking are available in most hypertext systems. The implications of using these links and their effect on usability have been previously discussed.

• Task applicability of links. Regardless of link type, the relevance and appropriateness of links for task performance could be an important consideration for hypertext system design. Intuitively, it seems the establishment of all links must be task-related to enhance performance with the system. However, rules for determining priorities and relevance of nodes in a network are not well established

Network configuration. Network configuration refers to the variety of system arrangements (configurations) for the nodes and links. Horton, Soderston, Garstka, and Park (1988) have provided examples of such network configurations (with no regard for whether the linking is at an organizational or relational level): (1) webbed network, (2) hierarchy, (3) expandable/collapsible (as in Guide™), and other formats. Conceivably, the network structure can have an impact on system usability, especially if improper configurations are utilized. For instance, expandable/collapsible configurations may be useful to hide unnecessary information from those who have no need for such. A webbed network may not be ideal for such an information management objective.

Locational cues. Locational cues are specific tools or techniques used to provide users information regarding their location in a hypertext system. Many authors have indicated the potential for disorientation in a large maze of nodes, especially if large numbers of relational links are established and heavily utilized (Conklin, 1987; Jones, 1987; Vail, 1989). Others (Akscyn, McCracken, and Yoder,1988) have held that less flexible network designs, such as strict hierarchies, obviate the need for locational aids; however, artificially imposed limitations do not take full advantage of hypertext capabilities.

A variety of locational and navigational aids has been proposed and implemented. Graphical browsers and maps, path tracing mechanisms (path history), indices, and dynamic tables of contents are only a few of the many aids used in hypertext systems. Monk (1988) investigated two versions of a hypertext system for time to answer questions regarding a computer program. One version was equipped with an overview map while the other was not; both versions were compared to a simple scrolling text file. The hypertext system without the map performed significantly worse than the scrolling text, while no significant differences were reported between the map version and scrolling text. Hammond (1989; cited by Nielsen, 1989) investigated the influence of graphical maps on a variety of dependent measures. One system was augmented with an overview map while the other system made no use of a map. The researcher compared the two versions for browsing through a hypertext system to retrieve information; dependent variables were answer accuracy, task completion time, and the ratio of newly visited nodes to previously visited nodes (new-to-old ratio). The third dependent variable is assumed to be a measure of navigational difficulty since if users are finding navigation problematic they are more likely to return to previously visited (old) nodes to reorient themselves. There were no observed significant differences between the systems for answer accuracy and task completion time; however, the map-augmented system performed significantly better for the new-to-old ratios, suggesting the value of maps over no navigational aids.

Unfortunately, the specific features required to create good maps are not easily definable or quantifiable. Hence, graphical maps and browsers themselves often vary based on idiosyncratic features. The same holds true of other navigational aids such as indices or tables of contents. There is some empirical support (and intuitive speculation) that navigational aids improve hypertext usability; however, questions still remain concerning the relative merits of various aids and objective methods for identifying the performance-enhancing features thereof.

Layout and lexical level issues. A variety of layout and lexical level considerations may influence hypertext usability as well. In general, these lexical issues are more easily quantifiable compared to the more nebulous topics of system design considered to this point. Most of the important issues are individually listed and discussed below.

- Screen size. Reisel and Shneiderman (1987) report a 50% decrease in task completion time for programmers answering questions using a 120-line display as compared to a 22-line display. A 35% decrease in task time was reported for a 60-line display as compared to 22 lines. Additionally, Bowers (1990) showed that a large 25-in. display performed significantly better than a smaller 11-in. display for multi-window tasks. Large screens may be beneficial as well for information retrieval tasks in hypertext environments. Many hypertext systems associate each node with an individual window (a one-to-one mapping); hence, the large screen may be especially advantageous for multi-tasking situations where users are accessing multiple nodes.

- Link indication. Link indication refers to the coding mechanism or technique used to indicate the presence of an active link such as bolding, surrounding, or otherwise demarking the linked item. Additionally indications are used to delineate the referent area (e.g., a portion of text) in the destination node. Many researchers (e.g., Raskin, 1987; Shneiderman, 1988) have discussed the issue of link indication. Initially, it is important to distinguish between embedded and explicit menus. An explicit menu provides link choices in an area of the node which is physically removed from the main information (be it text, graphics, or other formats). The KMS system employs such a scheme (Figure 11). Hence, due to their physical locations, explicit menu items require little coding. The coding challenge in these systems is to indicate the referent information in the destination nodes once the link has been traversed.

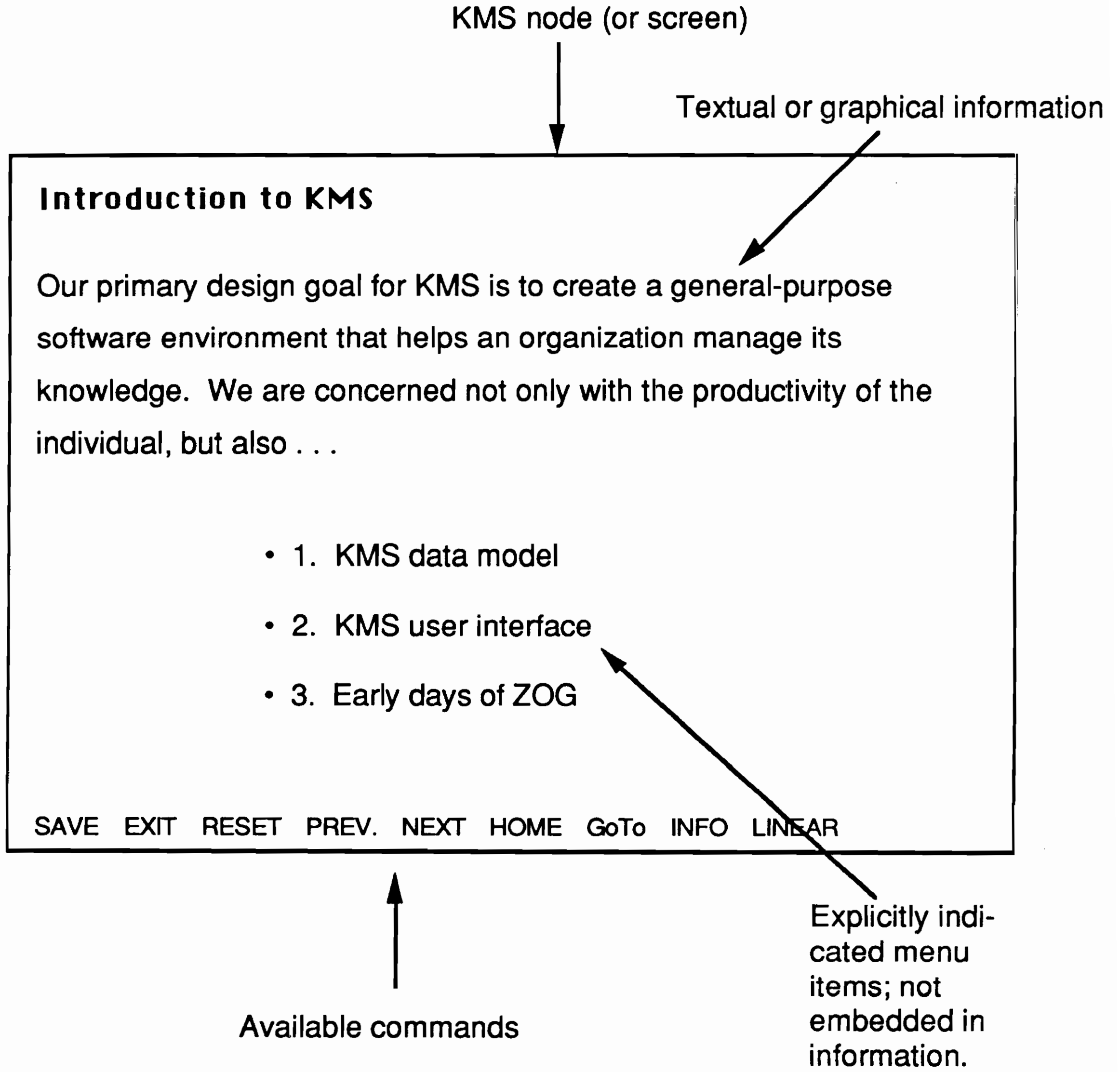


Figure 11. Sample screen from the KMS hypertext system. The indication of links is explicit since available links are physically separated from the information (similar to a menu) in the node. (Adapted from Akscyn, McCracken, and Yoder, 1988).

KMS uses no specific scheme; however, other systems use temporary highlighting of material, surrounding information with boundary lines, etc. The use of coding schemes may depend on whether the referent information is a single word, a region, a picture or graphic, or perhaps a sound or auditory message (in a hypermedia system).

Conversely, embedded menus embed the linked word in the information. Thus, in a graphical world map embedded menu, the user would view a world map and would simply click on the country of interest to access pertinent information. No other methods for accessing the information exist. In a textual node, the linked word is simply part of the material in the nodes but is somehow indicated as being a linked ("hot") item. Many techniques have been used for such indication; for instance, HyperTies uses word highlighting as the mechanism for indicating a link (illustrated previously in Figure 8), as does the Guide™ hypertext system. Guide uses a secondary coding scheme of dynamic cursor shapes. The shape of the cursor alters as it moves within the boundaries of a highlighted (linked) word; the cursor shape indicates the link type. Hence, a plus sign ("+") may indicate a link to more text while a cross ("x") may suggest a link to graphics. The Texas Instruments HyperTrans system surrounds linked words with carots (e.g., <linked word>). No specific coding scheme has been universally applied. The use of a variety of coding schemes, such as color, shapes, and special symbols is conceivable; however, the acceptable coding scheme may be application specific. For instance, the use of symbols may be inappropriate for domains which, by nature, use a multitude of symbols and notations, such as physics or statistics. Additionally, it is the intention to minimize interference with normal reading and operation. Coding unobtrusiveness must be a goal of link indication schemes.

Finally, embedded menus must also grapple with the problem of indicating the referent information in the destination node. No specific scheme has emerged as the optimum. Again, such coding may be task specific.

- Data density. Data density refers to the amount of information in a single node. Measures of density may include length of nodes, or amount of information in terms of number of characters or number of different topics addressed in a single node. Shneiderman and Versteeg (reported by Shneiderman, 1987) compared information densities of 9, 18, and 34 lines of information per node. Subjects were required to perform identical tasks using each of the systems. Mean task times for the 9, 18, and 34 line systems were 634 s, 593 s, and 513 s, respectively. The authors report these values were not significantly different, yet a definite trend is observable: as number of lines increases task time decreases. These results may be intuitively satisfying since a system with higher information density (more lines per node) should have fewer nodes than one with lower density. Users are not required to traverse as many nodes and are perhaps better able to maintain control of their movement through the system. For most systems the optimum information density may depend on the parsing techniques used and the structure and requirements of the information domain.

User Variables Influencing Usage of Hypertext Systems

The final group of variables effecting hypertext usability concern the user. There are a multitude of individual differences amongst users, many of which could potentially impact system use.

User domain knowledge. Domain knowledge refers to users' familiarity with the information on a variety of levels:

1. familiarity with inter-relationships within the information,

2. familiarity with significance or importance of particular issues of the domain, and

3. knowledge of the relationships between the system's information and other information.

In essence, users may have varying levels of knowledge regarding the information contained in a particular database. Conceivably, such a domain knowledge dimension exists for most information domains, anchored by complete novices at the low end and subject matter experts (SMEs) on the high end. Users can fall anywhere in the knowledge continuum. For example, an experienced biologist and a first-year biology student have quite different levels of knowledge concerning the topic of biology; this difference in domain specific knowledge could play a role in using a biology database, for instance (Egan, 1988). It is not possible to identify the multitude of levels for this variable.

Knowledge is not easily quantified unless standard tests with high reliability and validity are designed to pool the level of one's knowledge (e.g., the GRE). However, there is no doubt that people have differing levels of knowledge about various topics. This knowledge may be gained through experience, interest, education, and other mechanisms. Egan (1988) indicates that evidence (in the form of empirical data) is scarce concerning the specific influence of knowledge on users' success in system usage, but that it is reasonable to hypothesize such knowledge could impact system use.

User system knowledge . System knowledge refers to users' familiarity with the host system, such as:

1. familiarity with system operations in task performance, and

2. knowledge of available system functions and their utility.

In general, system knowledge is the level of user familiarity with the particular software (and hardware) used to present and manage information, regardless of users' familiarity with the information (domain) itself. For instance a user may be familiar with all the functional capabilities of Word Star™ (a word processing package), but have no knowledge concerning the subject matter of an essay written in Word Star. While some information retrieval systems are designed for public access and require little or no specific system knowledge to operate, others are designed specifically for a particular information domain and the appropriate user population. These systems are, perhaps, more sophisticated and include more functionality than public access systems (as would be available in a shopping mall or airport). As with domain knowledge, a user population may have many levels of system knowledge based on experience and frequency of usage. Egan and Gomez (1985), for instance, found system experience to be an influential factor by accident. These researchers had anticipated that subjects' typing skill would predict performance on a text editing task. However, typing skill began to predict performance only after subjects had considerable experience with the text editing system and its operations.

General level of computer expertise. This issue refers to users' competence with computers in general, which is usually determined by user experience (years of use, programmer vs. non-programmer, etc.). Elkerton and Williges (1984) found novice and expert computer users used significantly different search strategies in searching a hierarchical file-search environment. For instance, novices were more than twice as likely to page through the file (as in page scrolling) while experts were more likely to use string search capabilities; experts were also found to rarely use line scrolling mechanisms as compared to their novice counterparts, who used scrolling significantly more often. Perhaps user experience is influential in hypertext system operation; this effect may be especially pronounced for systems with a large number of features with which novices are quite unfamiliar.

Spatial ability and spatial orientation ability. Vicente, Hayes, and Williges (1987) and Egan, Bowers, and Gomez (1982) have shown a high degree of correlation between spatial ability and successful usage of computer systems. Vicente et al. illustrated this effect in retrieving information from a hierarchically organized database where users' spatial ability predicted 33% of performance variance, while Egan et al. demonstrated such effects with task performance in a text editor. Campagnoni and Ehrlich found a high correlation between spatial ability and task performance in a "hypertexted" help system. Task performance was measured by task completion time and number of returns to top level nodes, and correlations between spatial ability and each measure were $r = -0.75$ ($p < 0.005$) and $r = -0.65$ ($p < 0.01$), respectively.

Motivation for system use. Users' motivation in using the system may be a factor in hypertext usage as with using any system. Motivation levels may be correlated with the application and system setting.

Summary

Unfortunately, there is little empirical support for many of the discussed issues and variables. It is, nonetheless, true that hypertext and hypermedia are being implemented in a multitude of settings and applications with no regard for the influence of the discussed variables. It is particularly disturbing that, in the absence of scientific support for the utility of these systems, their popularity and use are growing. Hypertext systems are now becoming "fashionable" for information management tasks of all types, including information retrieval, education, and entertainment. Most potential designers, implementors, and users of hypertext seem unperturbed by the lack of empirical support for the utility of these environments. The duty of human factors engineers is to employ empirical methods (and tools) to evaluate the efficacy and viability of systems based on their facilitation of task performance or other measurable benefits. Such an approach has not been adopted in evaluating hypertext environments as compared to other alternatives. Hence, research is needed to test many of the discussed issues; in particular, the inherent utility of the hypertext approach should be questioned and its relative merits (or lack thereof) compared to other online systems should be investigated.

Rationale and Approach

Hypertext advocates claim the most salient advantages of these systems are the modularity of information, by parsing into nodes, and non-linear access to information through linking. Admittedly, both modularity and linking are attractive components of hypertext systems; however, the current research does not investigate the inherent utility of online information nor the advantages of parsing information. This study is concerned only with linking of material for information retrieval applications, which are considered to be a large area for the use of hypertext systems.

As defined earlier, there are two types of linking: organizational and relational. Prior to the proliferation of hyper-environments, organizational links (e.g., hierarchical or alphabetical linking) were used for information retrieval and presentation. Loosely speaking, most system (database) designers are familiar with hierarchies or alphabetical organizational structures and use them regularly. Conversely, the newly popularized relational links have been purported to be the essence of hypertext due to their relevance and flexibility (Conklin, 1987). Relational links are established based on some user- or author-derived dimension(s), which is (are) presumed to be beneficial for task performance. A significant amount of time and effort must be invested in creating relational links, especially since no universally applicable or acceptable quantitative (or automated) linking methods exist. In addition, system resources must be devoted to providing and maintaining relational links in addition to standard organizational linking.

In applying hypertext to an information retrieval scenario, the assumption is that the relational links established in the system may be ideal for accommodating retrieval tasks which specifically make use of relationships captured by the established links. However, from a human factors engineering standpoint, links are of no value if they do not enhance performance in some measurable fashion over other converntional systems. The issue of enhanced performance is especially questionable with current hypertext systems, which use author-established links and provide little or no user configurability. Many hypertext enthusiasts claim "ideal" hypertext systems offer the flexibility of user-derived links, thus obviating the restriction to author-established links. While this functionality is helpful for expert users, it is not a current option for commonly available systems since most users are not thoroughly familiar with the information domain, may not be willing to invest time in creating such links, or may be hesitant to establish links due to their lack of knowledge concerning system operations. Further, author-established links, especially relational links, are likely to fall short of fulfilling the needs and informational requirements of all retrieval instances, even for a single information domain. It may be simply impossible to predict all potentially useful relationships which could exist among the nodes. Hence, the question remains: Is the hypertext approach of non-linear access to information appropriate for information retrieval tasks? The experimental question was further decomposed into the following:

1. Do relational links, tailored specifically for task performance, provide a performance advantage over other linking structures and, if so, to what extent?

2. In situations where task requirements are not perfectly matched (or served) by relational links, will a negative effect result from including these links? If so, to what extent as compared to other linking configurations?

3. In a condition such as that described in 2, will organizational links suffice (as the "default" linking structure) or is the conventional linear structure more appropriate?

4. Are the extra cost and effort associated with incorporating relational links worth the gains in task performance? Perhaps the extra effort of relational linking is approaching a point of diminishing returns.

5. Do linear text and hierarchical linking perform as well as relational linking?

Additionally, the number of links and nodes a user must traverse to complete an information retrieval task may play an important role. For instance, an information retrieval problem requiring the traversal of two links may present no problems in any linking structure, but one requiring six traversals may be significantly more prone to error and more time consuming in certain linking structures. Hence, the questions outlined above must be considered under the mediating variable of number of required links.

Such fundamental questions of utility and usability of hypertext should be addressed before pondering the more molecular issues such as node size and link management. The hypertext approach, while currently popular, should be subjected to scientific investigation. Empirical results may suggest these systems are inappropriate for some applications and viable for others; it is hoped that the current study will lead to more judicious applications of hypertext systems.

METHOD

Online Database

An online geography database (called GEO) was created specifically for this experiment. The database was a HyperCard™-based system; all code was written in HyperTalk™ version 1.2.2. The reader is referred to Goodman (1988) for more detailed descriptions of the HyperCard environment.

System description. GEO contains 187 nodes providing detailed information about the six countries of North Africa: Algeria, Egypt, Libya, Morocco, Tunisia, and Western Sahara[1]. This particular region of the world was selected with the contention that very few, if any, potential subjects would have significant knowledge of the area. Detailed explanations of system linking configurations are not included here since the linking configuration was a variable of interest in the study. GEO's organization differed depending on the linking structure imposed. Graphical depictions and detailed descriptions of the system organizations are provided in discussing the independent variables.

GEO contained information pertaining to the demographics, economy, education, geography, and government of each country. The nodes provided

---

[1]As a political side note, Morocco, claiming territorial and religious rights, invaded the Western Sahara (W.S.) in the early 1980s and currently occupies a large portion of the nation. As of 1990 the United Nations had not resolved the question of W.S.'s independence (or lack thereof); thus, the GEO database considered W.S. an independent and sovereign state.

varying levels of detail regarding each of the topics. Each node was an individual entity whose information was meaningful in a stand-alone fashion. Yet, all nodes have relationships with other system nodes and provide richer information when these interrelationships are exploited . Again, the linking structures (to be discussed) determine the relationships available for use.

Information parsing. A number of geography/history texts, books, and articles were consulted; however, the bulk of information was extracted from the "Country Study" series of geography books published by the U.S. Government in conjunction with the American University. The "Country Study" book series devotes entire volumes to individual countries around the world. Experts in the region or country of interest contribute information concerning a myriad of topics for each country (or book). These topics include, but are not limited to, education, government, military, political history, social history, economics, transportation, demographics, physical geography, and social attitudes and customs. While no two books are identical, the topics included and their general organizations are quite similar. The author chose to draw considerably from the "Country Series" books since they are well respected as solid and detailed introductory texts (Nelson, 1979a, 1979b, 1986; Nyrop, 1983).

The "Country Study" books use a hierarchical organization of information. Books are divided into large chapters each of which contains smaller sections. Each section is comprised of sub-sections; these sub-sections often contain further subdivisions. An accepted method for data parsing is to use the original document as the guiding framework. Accordingly, in the creation of GEO, the section, sub-section, and sub-sub-section entities in the books were converted into individual nodes. A one-to-one conversion from linear to parsed was followed: each section or sub-section to be included in GEO was assigned an individual node. Figure 12 illustrates this conversion process. As seen in the figure, titles of created nodes were not differentiated by font, color, alpha-numeric coding combinations, or other mechanisms for signifying order or levels of specificity/generality. The absence of coding was established to avoid providing locational cues for certain linking configurations. Title coding schemes such as those used in the original document may have been appropriate for a hierarchical or linear organization; however, those codes would not be appropriate in a pure network configuration.

Each node was confined to a text area of 19 lines x 64 characters (Figure 13) to eliminate experimentally uncontrollable scrolling rates. Kolers, Duchnicky, and Ferguson (1981) found subjects prefer highly varying scrolling rates; this variability is so high that in a follow-up study, Duchnicky and Kolers (1983) allowed for user-adjusted scrolling rates in the range of 0.2 to 5 lines per second. At times information reduction and/or paraphrasing were required to properly position the original text within the confines of the newly created node. No significant alterations to the information resulted: all critical information was maintained and no false information was included. The physical appearance of system nodes was identical to that shown in Figure 13. Minor alterations to button labels or menu items resulted according to the linking structure. Figures 14, 15, 16, and 17 present sample nodes for the linear, hierarchical, network, and combination structures, respectively.

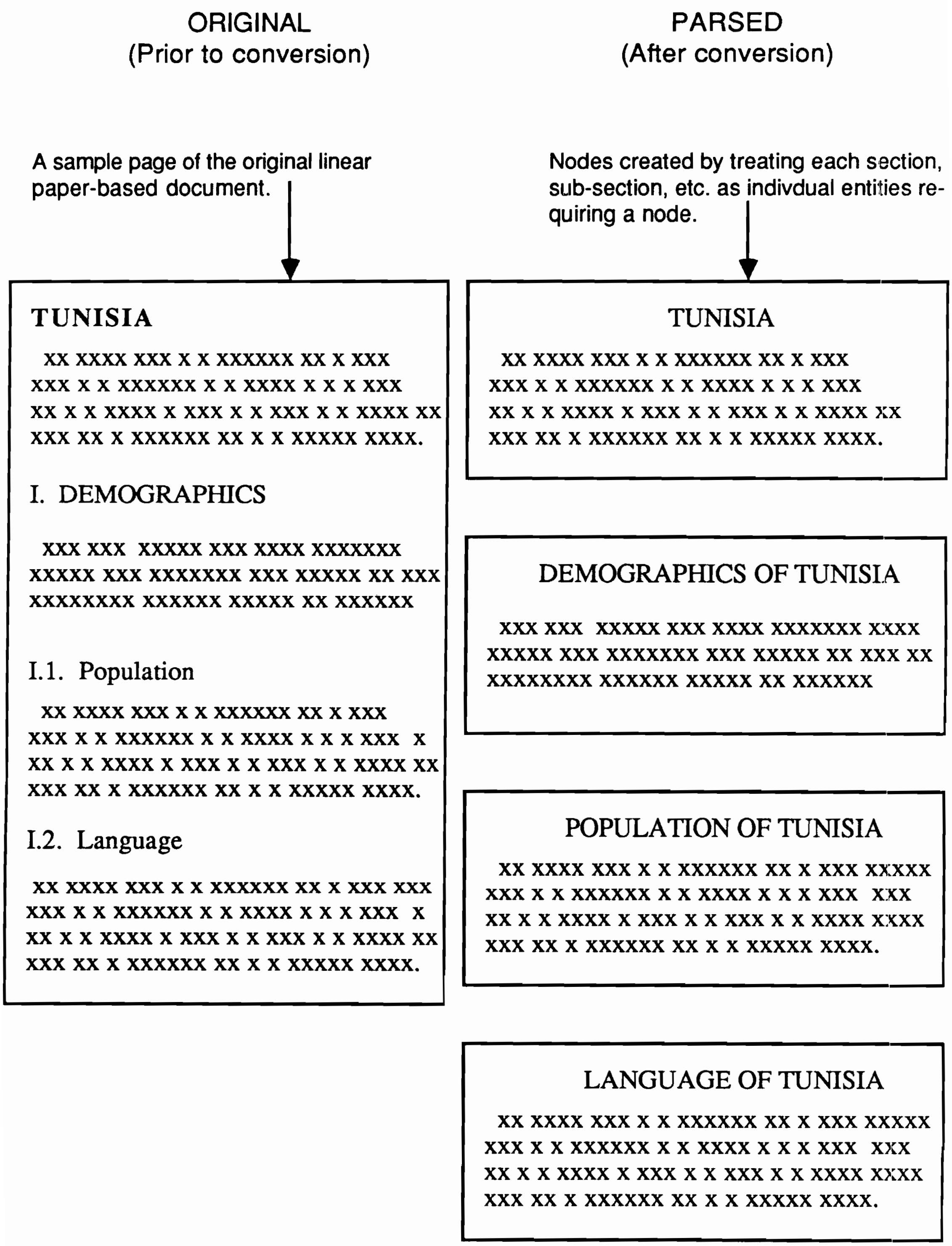


Figure 12. Converting paper-based material into individualized nodes.

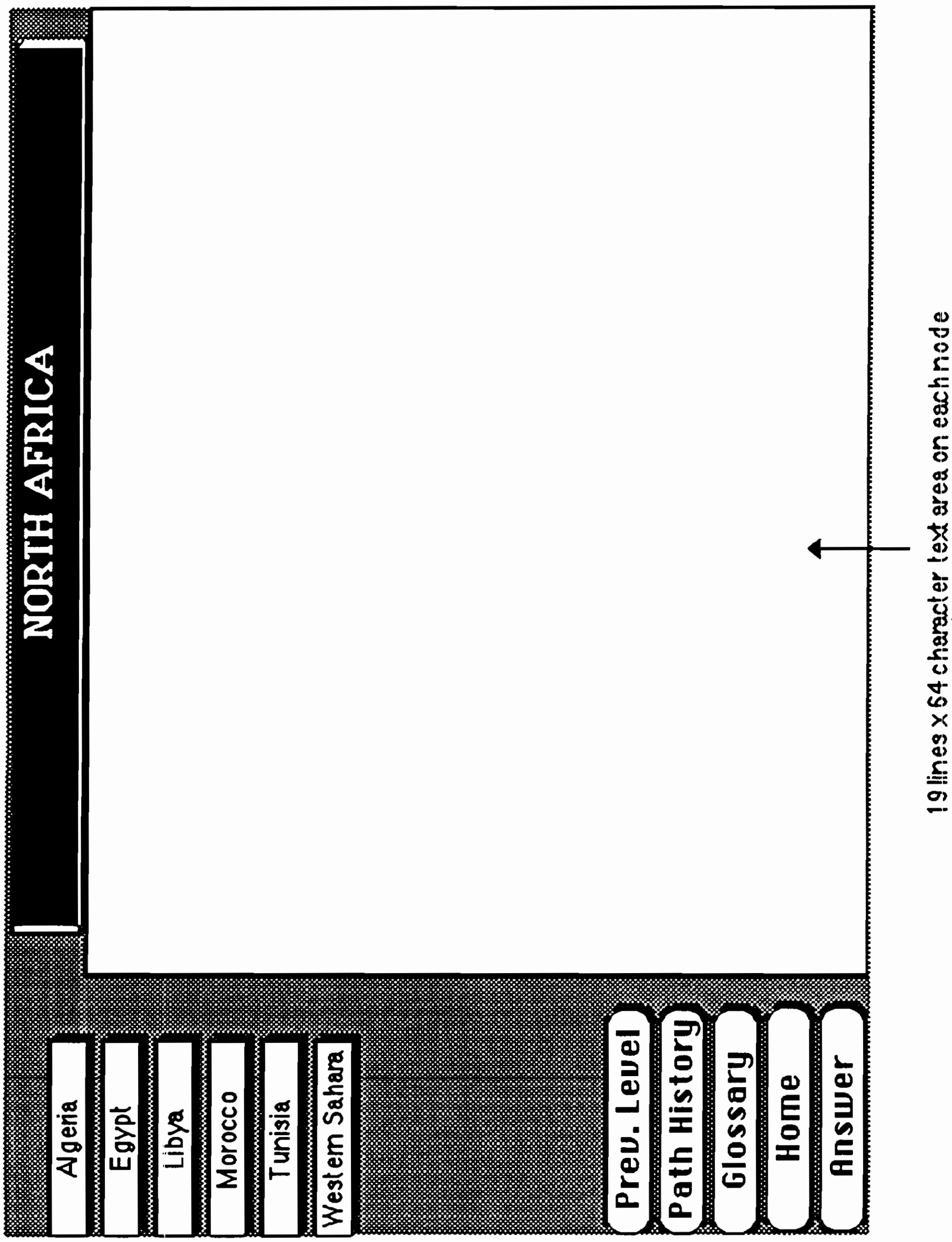


Figure 13. Allowable 64 X 19 text area in a GEO node.

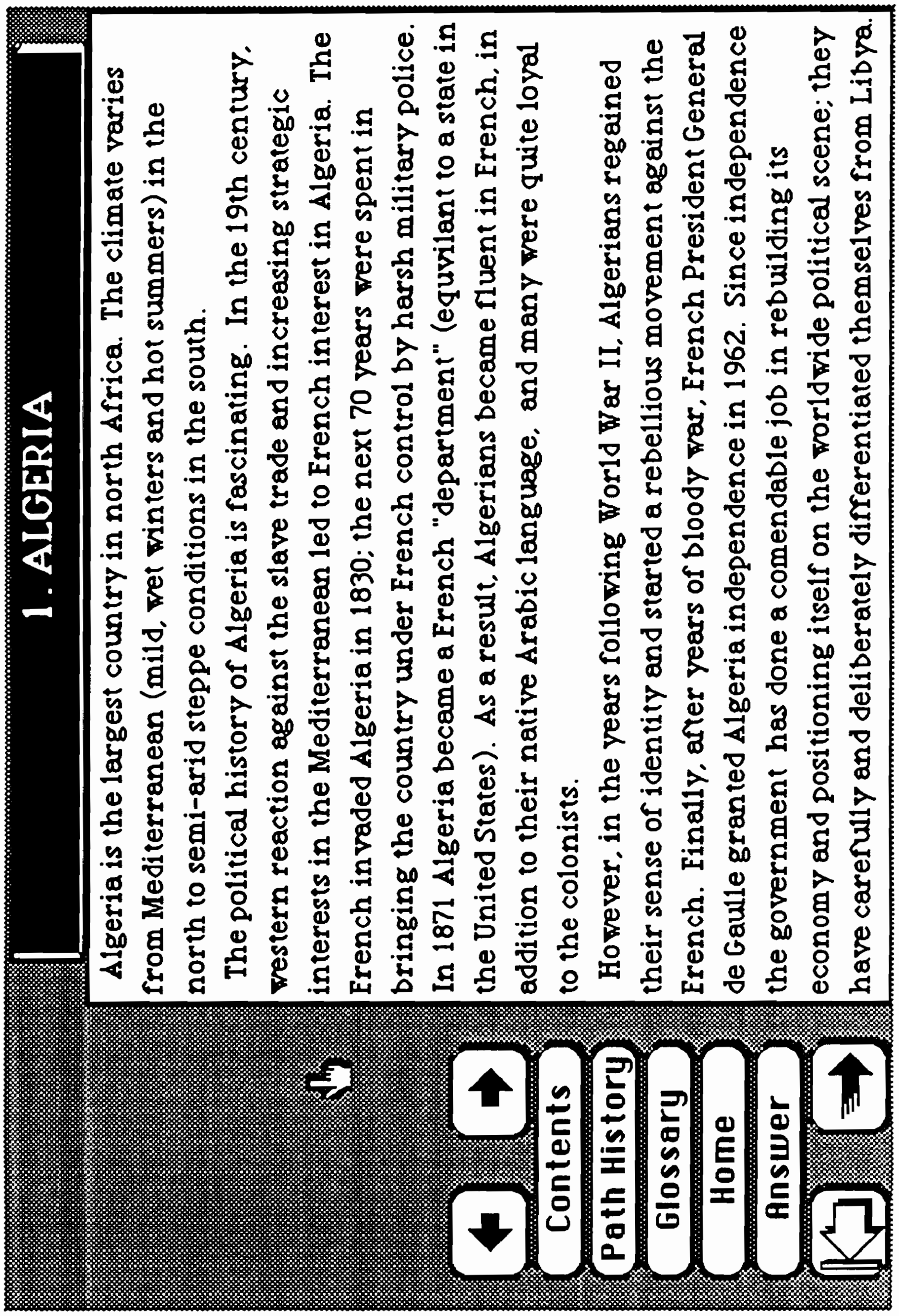


Figure 14. Sample node from GEO in a linear configuration.

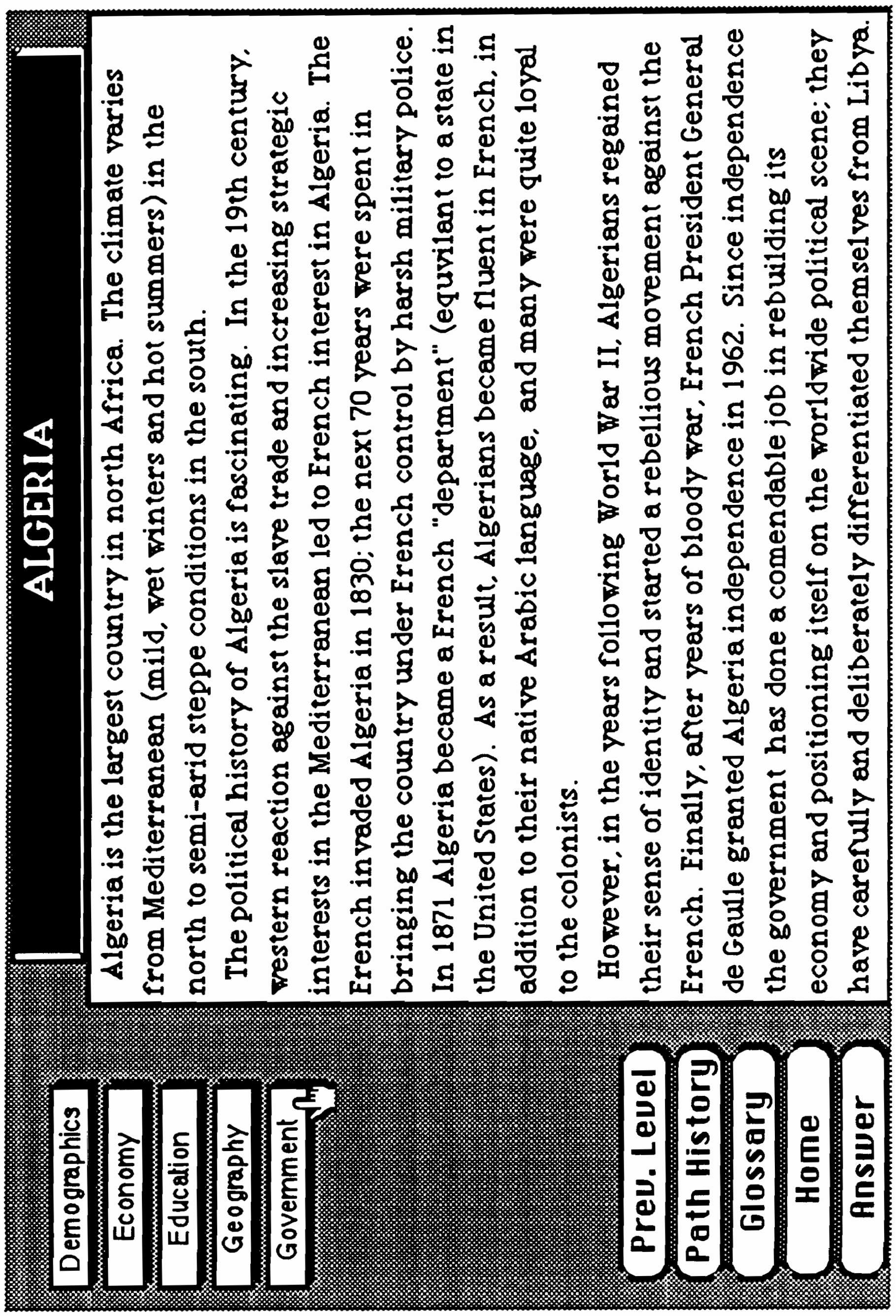


Figure 15. Sample node from GEO in a hierarchical configuration.

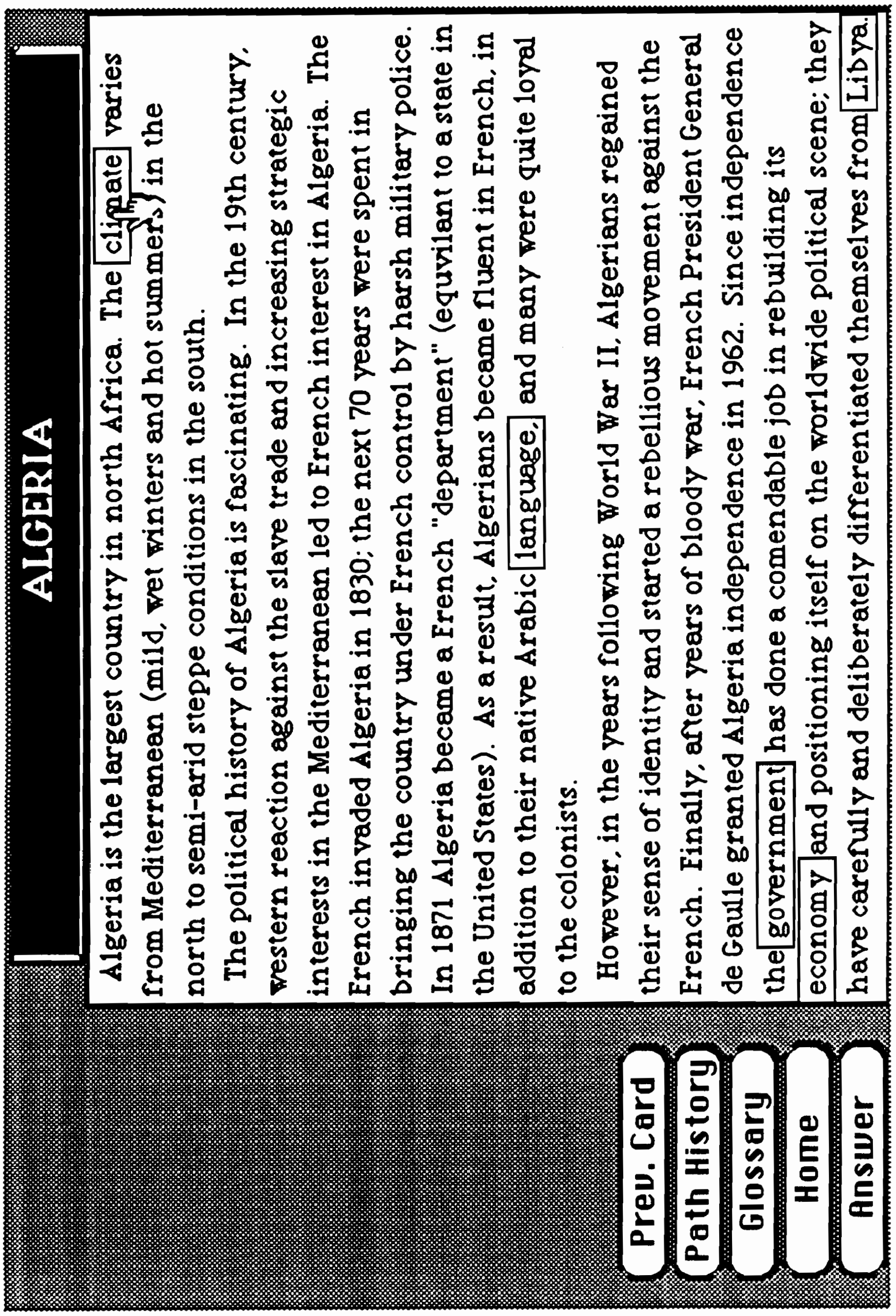


Figure 16. Sample node from GEO in a network configuration.

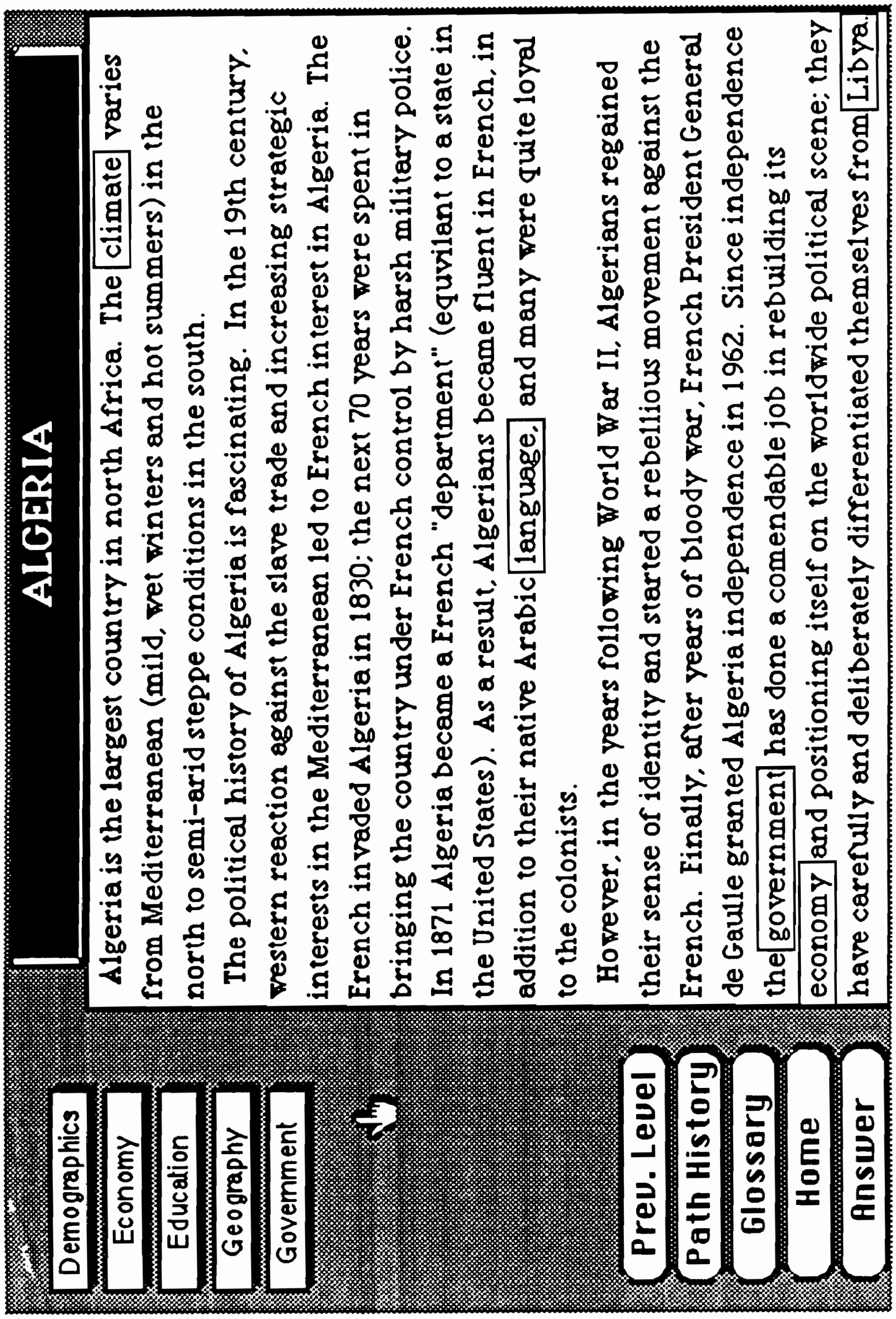


Figure 17. Sample node from GEO in a combination configuration.

Independent Variables

Three independent variables were investigated. The following sections provide detailed explanations and descriptions of these variables and their concomitant levels.

System linking structure. Linking structures are methods used to connect nodes of information in the network. Links are used to browse through system nodes in search of desired information. Browsing techniques are included to augment other search mechanisms such as string searches and query languages. Hypertext systems have been considered the optimal tool for browsing. This variable had four levels: linear, hierarchical, network, and combination linking structures.

- Linear. The linear structure simply connects nodes in a strict linear configuration, but maintains the hierarchical grouping of sections and sub-sections as found in the original books. Information was divided into six chapters (corresponding to the six countries of North Africa). Within each chapter information was organized as illustrated in Figure 18. As seen in the figure, chapters were comprised of five sections, each of which had five sub-sections. The number five was selected to maintain an acceptably low number of choices. It was decided that five choices, being on the low end of Miller's (1956) 7 ± 2, would not constitute a challenging number of options. This consideration was particularly important because the intention of the study was not to investigate optimal number of choices per node, especially since this number may

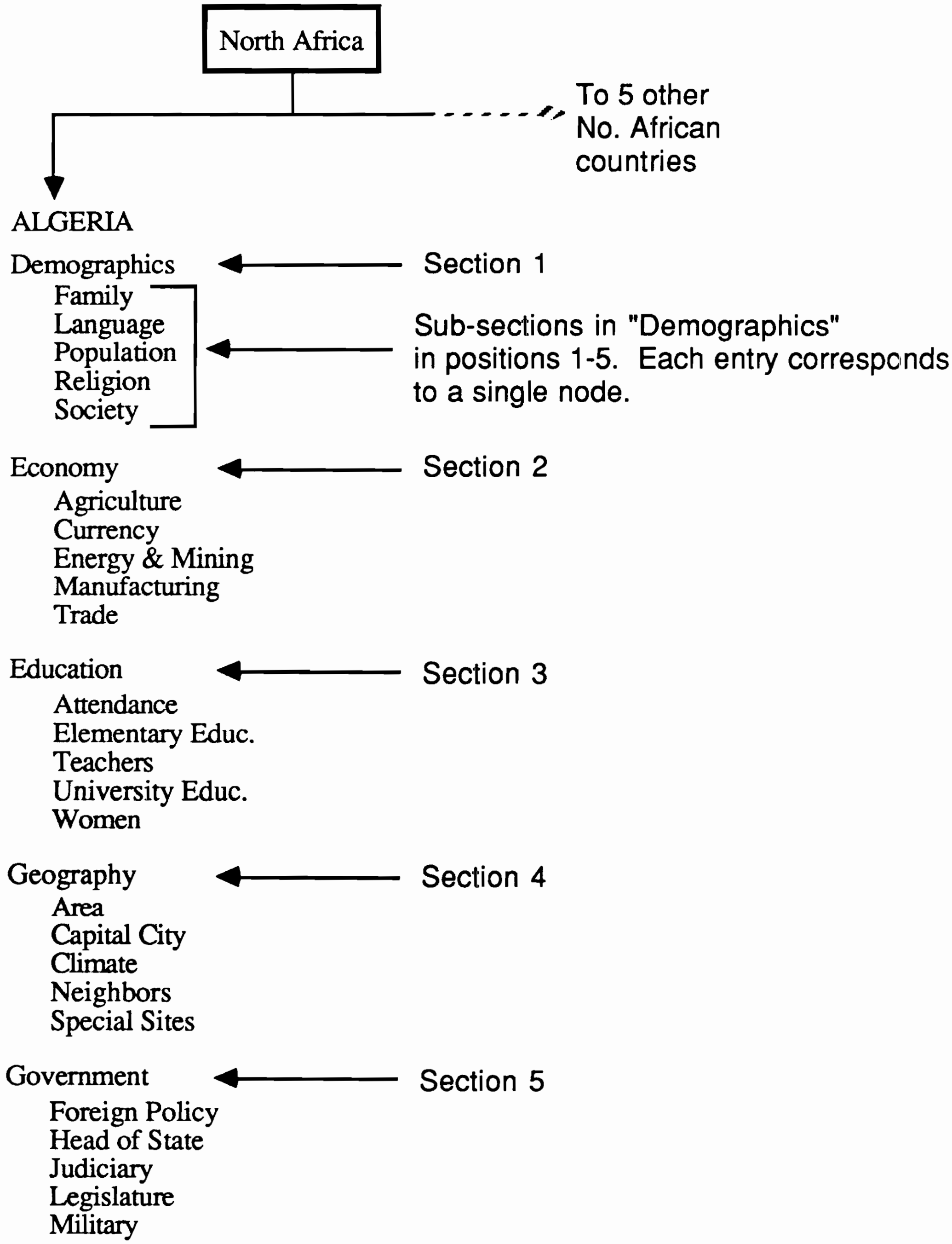


Figure 18. Information organization in the linear system structure. Each entry in the figure represents and individual node in the GEO system.

vary depending on the application and user experience. Rather, the linking structures were of interest, and it was hoped the low number of choices would reduce experimentally confounding confusion associated with an overwhelming number of options.

In the linear structure, users moved linearly from one node to another. For instance, in the Algeria chapter, to move from the agriculture node to the trade node one was required to pass through the energy, currency, and manufacturing nodes. Section heading nodes (labeled in Figure 18) could be accessed by one link from any node within the section or nodes in the immediately preceding section. These direct section access links allowed for more efficient navigation and were especially helpful for reaching the end nodes (e.g., the "Military" node) within a chapter. This type of navigation is typically found in books and traditional paper-based documentation. The minimum distance between any two nodes in the same section is defined as:

$$D = | i - j |, \qquad (3)$$

where i and j are the positions of the two sub-section nodes in the section; D is an integer whose units are links. The minimum distance between two nodes in different sections of a single chapter is defined as:

$$D = (|x - y|) + j, \qquad (4)$$

where x and y are the positions of the two sections and j is the position of the target node within its sections. Finally, the minimum distance between two nodes in different chapters is defined as:

$$D = (x + j) + 1, \qquad (5)$$

where x is the position of the section of the target node and j is the position of the node within the section. This equation is derived since the top node ("North Africa") can be accessed from any node in the system by traversing one link; from the top node to the target node requires $x + j$ links. The linear configuration was used as a control condition for task completion time, errors, and subjective measures.

- Hierarchical. The hierarchical linking structure maintained a nested relationship among the nodes as depicted in Figure 19. This structure provided a strict hierarchy as found in the "Country Study" books; however, the top ("North Africa") node was accessible from any location. As with the linear configuration, the organization of material within each country is identical. The essential difference between hierarchical linking and linear linking is the immediate availability of all options associated with a particular chapter or section, presumably allowing for more efficient movement within the document. The number of choices within a chapter or section was restricted to five for reasons discussed previously. Additionally, care was taken in system design to maintain the depth of the hierarchy at four levels, well within the 3 to 8 level range suggested by researchers of breadth-depth issues (Allen, 1983; Kiger, 1984; Miller,

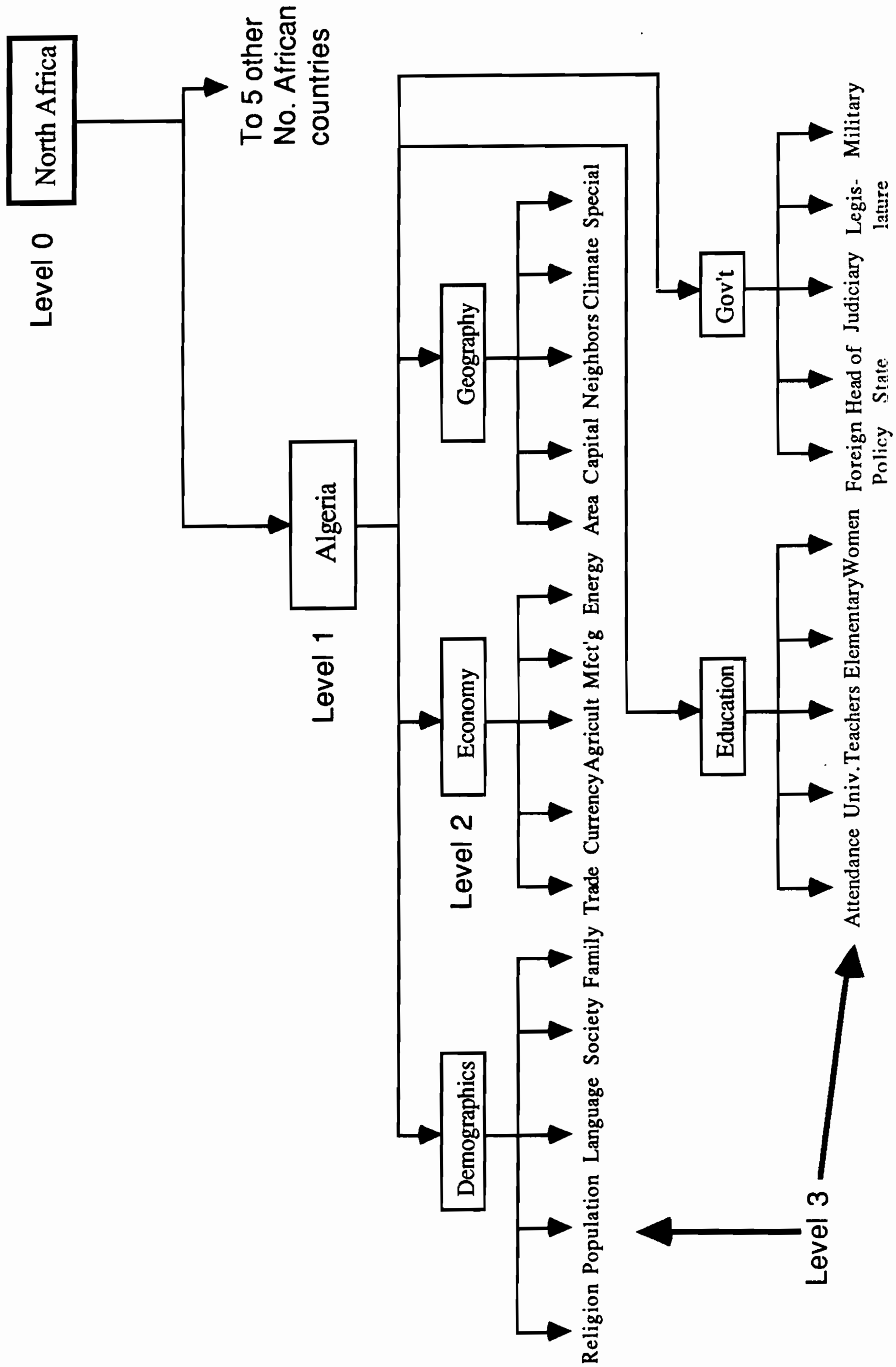


Figure 19. Information organization in the hierarchical system structure.

1981). Options within each chapter and section were organized alphabetically as determined by the menu panel design guideline (Figure 20) offered by Paap and Roske-Hofstrand (1988) . Assuming the top node to be level 0, the distance between any two nodes within the same chapter is predicted by formula (1). The distance between two nodes in different chapters is defined as:

$$D = j + 1, \qquad (6)$$

where j is the level of the target node. The top node can be accessed in one link from any locale; from the top to any node requires the number of links identical to its level in the hierarchy.

- Network. The term "network" has been applied in a multitude of settings and system configurations; however, the current study uses this term to refer to relational hypertext links as described earlier. The term "relational" is omitted to avoid confusion with relational databases, which are significantly different from text-intensive hypertext databases. Network links were established on the basis of keyword searches. The titles and subject matter of all system nodes were defined as keywords. Visual and text string scans located keywords in every node. The number of keywords located per node ranged between three and seven, with a mean of roughly four. To equalize the number of options per node in all system structures (to five), the author edited some nodes to bring the number of keywords per node to between four and six with most nodes containing five keywords. If considerable alterations were

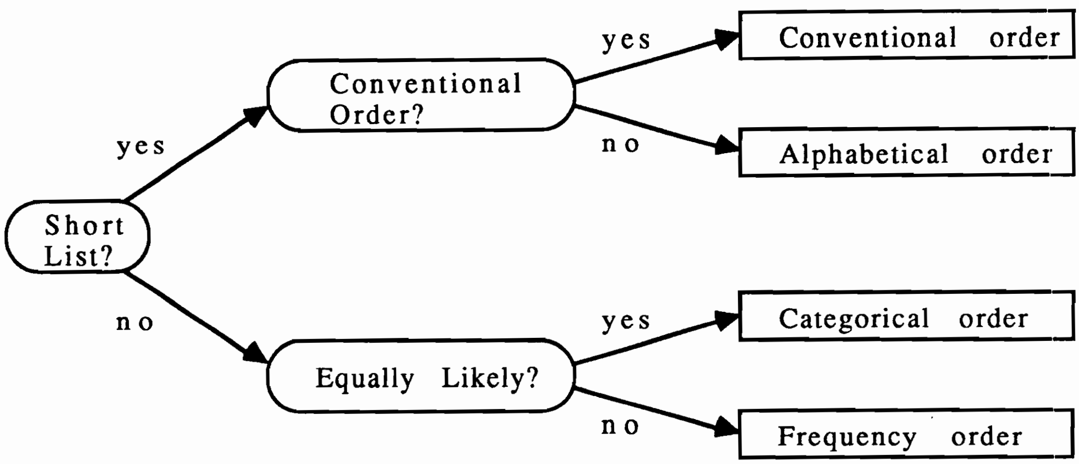


Figure 20. Paap and Roske-Hofstrand's menu design guideline. An alphabetical order was selected for the hierarchical and linear linking structures since there were few menu items with no conventional order. A conventional order was used in the network structure since the location and ordering of items were dependent on the text.

required the node was left unedited in order to remain true to the original documents; the number of keywords remained unchanged. Keywords referred to information contained in other system nodes. Hence, a link was established from the keyword to the referent node. A sample node in the network linking configuration is illustrated in Figure 21. Note the lack of regard for node position in hierarchical terms; network links are established solely on the basis of reference-referent relationships. Thus, links are established between nodes concerning the same country and those of differing countries.

In terms of the dimensions used to describe relational links (Figure 5), the created links were author-established, intra-document, one-to-one, internally meaningful connections. The dimension for their establishment was the reference-referent relationship as identified by the keyword searches. This type of network (relational) linking is typical of current hypertext structures. An arithmetic function for describing the distance between two specific nodes is possible for the network structure; however, the distance between any two randomly selected nodes cannot be predicted due to the lack of uniformity of network links. Hence, the distances between any two nodes may vary from one link to as many as six or seven links. As with all system structures the "North Africa" node is directly accessible from any node.

- Combination. This linking structure is simply the combination of the above described hierarchical and network structures. In effect, users have access to two types of linking and twice as many options (at most

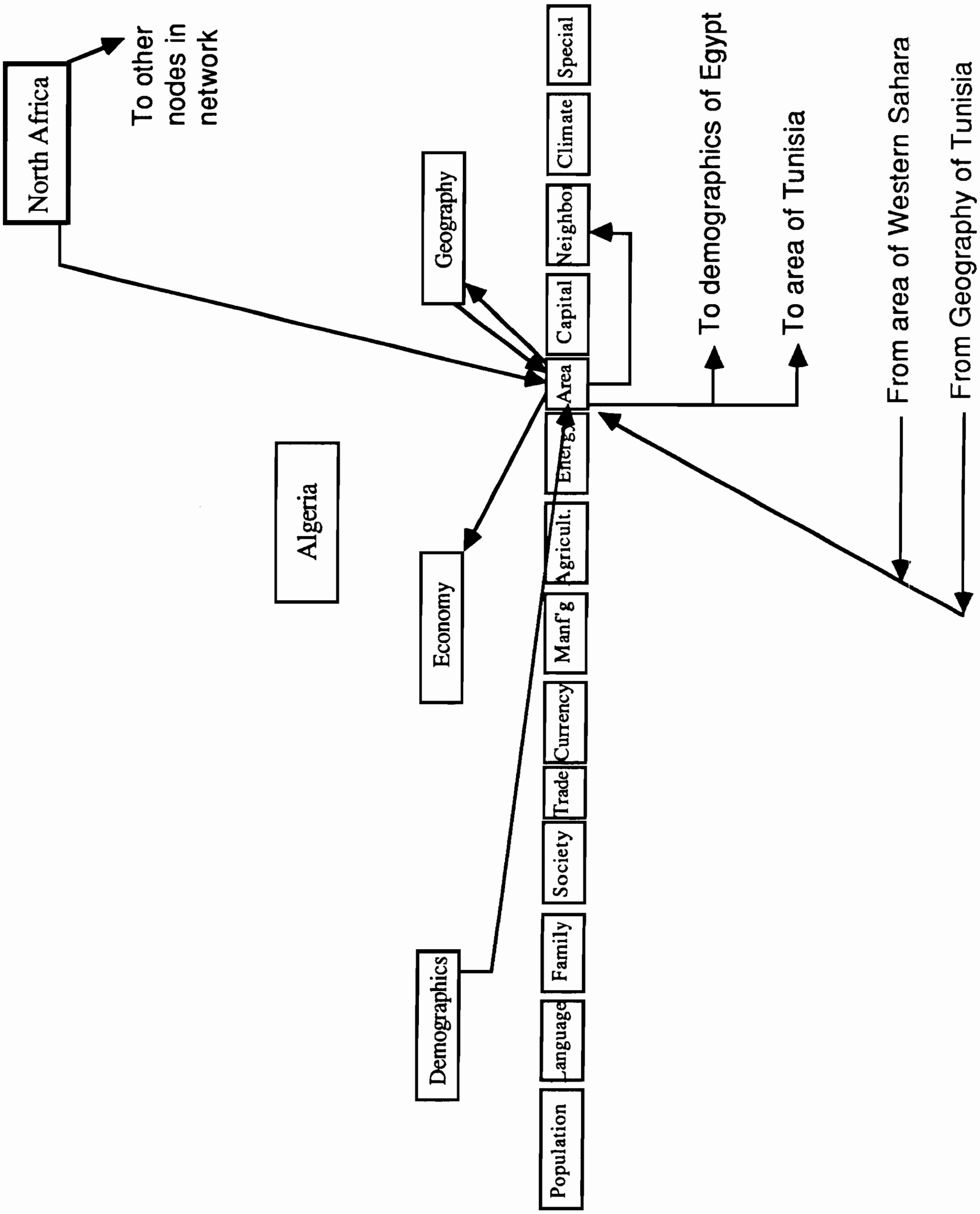


Figure 21. The "Area of Algeria" node in the network linking structure. Multiple links connect this node to various nodes throughout the document with no regard for their locations in hierarchical terms.

nodes) than for the hierarchical or network systems alone. Figure 22 illustrates the combination linking structure. Some hypertext experts (e.g., Conklin, 1987; Shneiderman, 1987) believe such a two-tiered configuration of organizational plus relational links is more beneficial than each in isolation.

Number of required links. The required links variable refers to the minimum number of links necessary to satisfy an information retrieval (IR) query using the most efficient route. There were three levels: 2-, 4-, or 6-link tasks. The number of traversed links is simply a description of the minimum number of required traversals; all subjects did not pursue this optimum path and, indeed, most used more than the minimum number in completing a task. Answers to 2-link tasks were provided by extracting information from one node; 4- and 6-link tasks required subjects to visit 2 separate nodes to properly satisfy the query.

Most modern databases provide sophisticated IR mechanisms such as string searching, query languages, boolean searching, and weighted searching. Belkin and Croft (1987), Dumais (1989), and Salton and McGill (1985) provided detailed discussions of IR techniques. In hypertext systems, one or more of these techniques is used in addition to browsing, which is the technique of interest in this study. Thus, in an applied situation, the supplementary search mechanisms would be used to narrow the focus of the search and bring the user closer to the desired information. Since quality of design of the search mechanisms and the user's ability to properly utilize the features will determine its effectiveness, it is reasonable to assume that an exact match will not always be found, resulting in the user being positioned in a node *near* the answer.

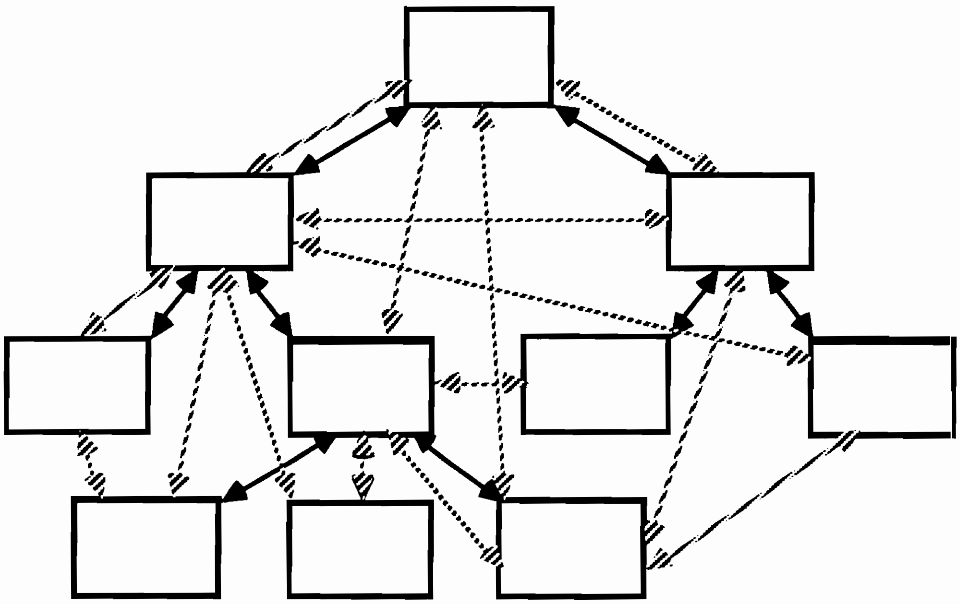


Figure 22. Example of combination linking structure.

Additionally, the system may be able to access only the user's general area of interest rather than the exact node containing the answer; moreover, many users prefer the browsing mechanism (Brown, 1988). Therefore, the 2-, 4-, and 6-link levels are of interest since they constitute a fairly representative range. Traversing many more than six links for completing an IR query was considered unrealistic given the available high-powered string searching and query language capabilities, which help narrow the focus of a search.

Task type. Task type refers to database programmers' judgements of whether IR tasks were best performed with the hierarchical or network linking structures. Prior to designing IR tasks it was hypothesized that, due to the inherent differences between the hierarchical and network systems in available links at any given node, certain tasks may be logically suited to the hierarchical structure while others may be better suited to the network structure. Intuitively this hypothesis seems reasonable; however, no known criteria were available for effectively dichotomizing IR tasks. Further, it seemed such categorizations would be subjective decisions based on interactions with the systems. Therefore, a pilot study was conducted using 12 database programmers (from the VPI & SU Computer Science Department) as expert judges. The database programmers were experienced users and creators of databases; all were familiar with hypertext. Their judgements and comments were solicited and assumed to reflect their technical understanding of the issues involved in IR. The experimenter constructed 12 IR tasks: three groups of four tasks requiring either 2, 4, or 6 links on the hierarchical and network structures. In addition, due to the flexible linking offered in the network structure, six tasks were constructed which required half as many links in the network structure as in the hierarchical (three groups of two tasks requiring 2, 4, or 6 links in the hierarchical system but requiring 1, 2, or 3 links, respectively, in the network system). Network linking flexibility in this context refers to the lack of regard for node position in the organizational sense; for instance, Figure 23 illustrates that moving from one node to another may require four links in a hierarchical configuration, but only two links in a network structure. The network links' ability to transcend hierarchical levels is an undeniable characteristic of network linking and was included in this investigation. Only the hierarchical and network structures were of interest in the pilot study since the combination structure was defined by these two and the linear structure was a control condition for which number of links (as defined by the hierarchical or network structures) was not meaningful.

Pilot subjects performed the 18 tasks, in random order, on both the hierarchical and network structures; counterbalancing procedures eliminated system order effects. After performing each task, subjects provided a rating of the difficulty to perform the task on each system structure and indicated which of the systems was better suited for performing the IR task. The ratings were performed on a 9-point modified Likert scale, while the system preference was performed on a forced-choice basis. The intention was to identify tasks judged reliably to be best suited to one system structure while not being significantly different on the two structures in terms of difficulty to perform. Hence, a binomial test was used to examine the statistical significance of the preferences, while an ANOVA was used to test differences in difficulty ratings for the two structures; a brief discussion is included here. Table 2 contains the results of the preference choices and difficulty ratings. Of the 18 tasks 13 were judged to be more

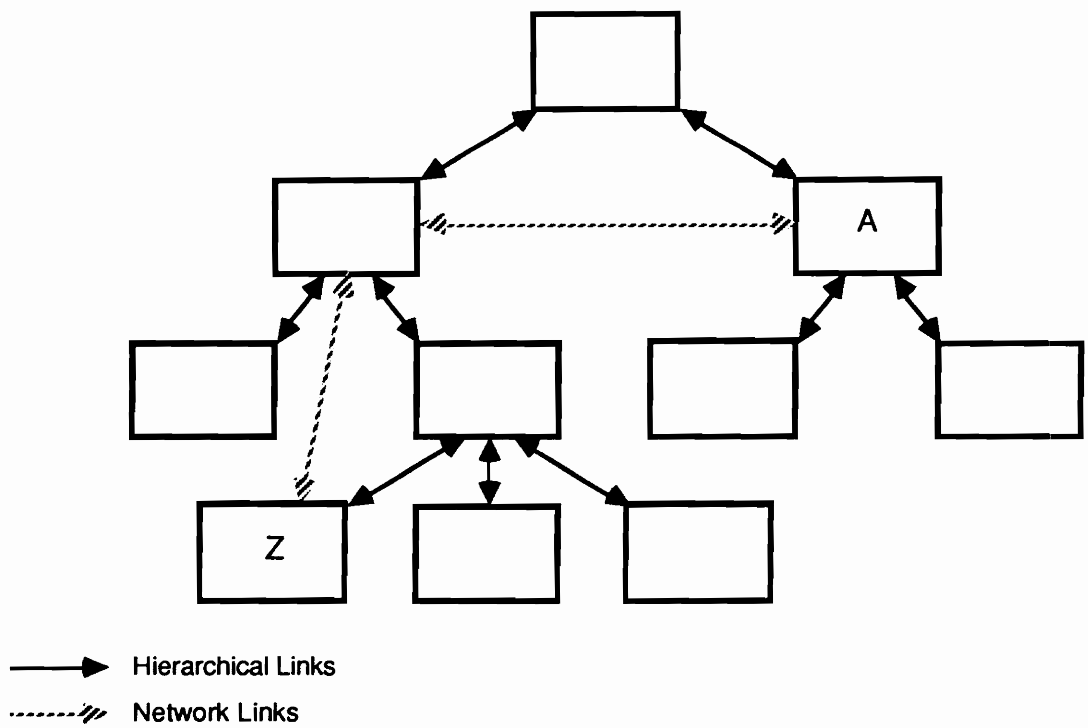


Figure 23. in the illustrated configuration moving from node A to Z requires two links in the network configuration but four links in the hierarchy.

TABLE 2. Summary of pilot data. "System Preference" refers to subjects' responses in dichotomizing tasks between the Hierarchical and Network systems, while "Difficulty" is simply the difficulty ratings assigned to each system per task. In both instances the preferred/easier system is listed, unless no significant difference was uncovered.

TASK	SYSTEM PREFERENCE	DIFFICULTY
2-Link Tasks		
1	Hierarchical	No difference
2	Network	Network
3	No difference	Hierarchical
4	Network	No difference
4-Link Tasks		
5	Hierarchical	Hierarchical
6	Hierarchical	No difference
7	No difference	No difference
8	Network	No difference
6-Link Tasks		
9	Hierarchical	No difference
10	Network	No difference
11	No difference	No difference
12	Hierarchical	No difference
2-Link Hier./1-Link Net.		
13	No difference	No difference
14	Network	No difference
4-Link Hier./2-Link Net.		
15	Network	No difference
16	Network	Network
6-Link Hier./3-Link Net.		
17	Network	No difference
18	Network	Network

suitable on one structure or another (as determined by a binomial test), and 13 failed to yield significant differences in difficulty (as determined by a Newman-Keuls post-hoc test, overall $\alpha = 0.05$). Task types had been identified, as presented in the "System Preference" column of Table 2; however, the goal was to include such tasks which failed to yield significant difficulty differences. Therefore, Table 2 was scanned to identify such tasks. Tasks 1, 4, 6, 8, 9, 10, 12, 14, 15, and 17 met the requirements. For 6-link tasks, two hierarchically suited tasks qualified (tasks 9 and 12); thus, the experimenter randomly selected one to be included in the study. The nine final IR tasks are included in Appendix I.

Analysis of subjects' comments indicated they used two important dimensions for dichotomizing IR tasks: (1) perceived effort in terms of number of traversed links, and (2) applicability (or utility) of available links to the IR task. Information gleaned from post-task comments indicates expert judges were more likely to base their decisions on perceived effort. Fewer links to task completion were usually considered a significant advantage. Since pilot subjects did not always follow the optimum path the perception of required effort was different for each subject. However, given equal perceived effort, subjects resorted to their perception of the appropriateness and applicability of the links they encountered during task performance. Determination of appropriateness relied on the level of agreement (or suitability) between users' understanding of task informational requirements and the available links. A combination of both criteria yielded highly favorable responses, as indicated by tasks 14, 15, 16, 17, and 18 being judged more suited to the network configuration.

The resulting task types were labeled hierarchical tasks and relational tasks. Since suitability of available links was an essential factor in the dichotomization, tasks judged best for the hierarchical system were believed to take advantage of the hierarchical links, hence the name "hierarchical tasks." The second group was assumed to make use of the relationships captured by the network links, hence the label "relational tasks."

Experimental Design - Experiment A

A three-way mixed factors factorial design was used for data collection (Figure 24): 4 linking structures X 3 levels of required links X 2 levels of task type. Linking structure was a between-subjects variable due to the threat of transfer of experience across the various structures. Number of required links and task type were within-subjects variables. The six IR tasks used in this portion of the study were defined by the combinations of three levels of required links and two categories of task type. For instance, task 1 was a 2-link relational task, while task 4 was a 4-link hierarchical task. Subjects performed all tasks on their respective systems with no knowledge of the task defining variables.

Experimental Design - Experiment B

A two-way mixed factors factorial design was used for data collection: 4 linking structures X 3 levels of required links. Task type was not relevant here since all tasks were judged to be best suited to the network linking structure (relational tasks). Data for experiments A and B were collected during the same session. The order of task performance was completely randomized; however, tasks 3, 6, and 9 were treated as belonging to Experiment B a priori. Figure 25

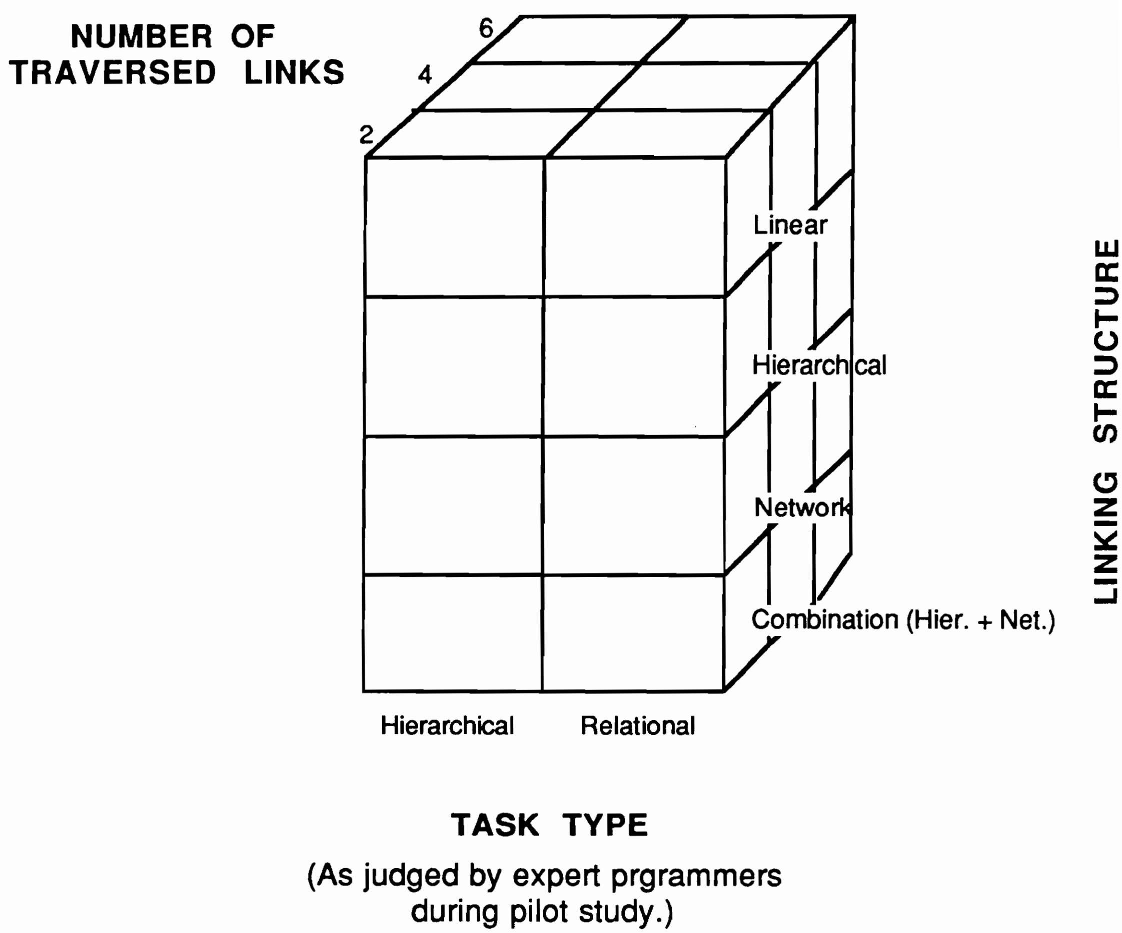


Figure 24. Experimental design for Experiment A, where number of required links to complete tasks is equal for both the hierarchical and network linking structures.

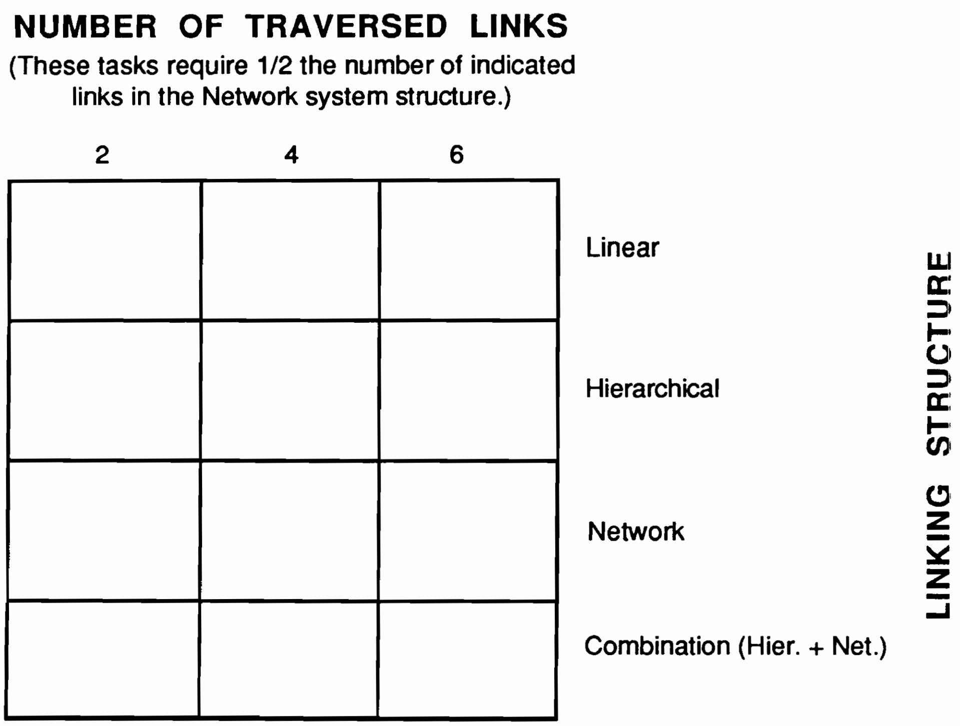


Figure 25. Experimental design for Experiment B.

illustrates this experimental design. The intention of this second experimental design was to investigate the performance of various linking structures under circumstances heavily favoring the network linking structure. Experiment B would uncover the advantages of the network configuration (if any) and determine the magnitude of these advantages as compared to other systems.

Apparatus

Database. The GEO database, described earlier, served as the database subjects used to complete their IR tasks.

Equipment . An Apple Macintosh II equipped with a 20-in. Sigma Designs SilverView™ monochrome monitor, an extended keyboard, and a mouse constituted the host system. All data collection was performed online.

Procedure

The 64 subjects were randomly assigned to one of four system linking structures (eight males and four females per structure). All subjects read a brief description of the study and signed an informed consent form (Appendix II). Participants also completed a questionnaire to determine their level of computer expertise and test their knowledge of North Africa. The questionnaire topics pertaining to North Africa, included in Appendix III, were considered basic facts, roughly 50% of which could be correctly answered by slightly knowledgeable persons. For example, a Middle-eastern student correctly answered 87.5% of the questions (7 of 8) and two American students with a casual interest in world geography answered 75% and 50% of the questions, respectively (none of these students participated in the experiment). Any subject correctly answering more than 25% of these questions was excluded from participation. On average subjects participating in the study correctly answered 4% of the questions (0 to 1 question of 8); most could answer none. Hence, all participants were considered novices in the information domain.

Training. Training materials were prepared for the subjects; the materials for each system structure are provided in Appendix IV. Participants learned about system linking structures, the functionality of different available buttons, and the use of glossaries and/or table of contents. Subsequent to reading the training materials, subjects performed three practice IR tasks. The experimenter monitored subject performance and provided guidance and feedback during and after task performance for the initial two tasks. The final practice task was performed without interference from the experimenter. All subjects felt comfortable in proceeding with data collection following the training period. On average training required 10-13 min. Upon completion of training subjects were encouraged to informally browse the system to further familiarize themselves with the system structure.

Data collection. Subjects completed nine IR tasks during the experiment (Appendix I). The order of task presentation was completely random. Originally, all tasks were presented online; subjects were assumed to read and understand the task and proceed with retrieving the necessary information. However, during the initial phases of the pilot study it became apparent that subjects, while completely understanding the task, preferred to double- or triple-check the task wording before committing to an answer. HyperCard 1.2.2

provides a usable screen area of 346 X 542 pixels, all of which was in use. Providing the IR task in a "pop-up" window was considered; however, any windowing would have obscured portions of the text, requiring constant alternation between the window and the node. Therefore, all IR tasks were transfered to paper. Participants received a strip of paper with the appropriate task, which they placed in a visible location and referred to at will. Upon reading and understanding the question, subjects clicked on a "Start" button to access the database. Task time was initiated at this point. Subjects navigated through the system in search of information to satisfy the query. Once they found the proper answer (or one believed to be the answer), subjects directly accessed an online answer sheet in which they typed their answer (fill-in the blank format). Task completion time was terminated upon accessing the answer sheet; this termination point was selected to eliminate non-relevant variability resulting from different typing speeds or differing formats for entering answers (e.g., US $ 1000 vs. one thousand U.S. dollars). Subjects completed a questionnaire (Appendix V) after performing all nine IR tasks. Finally, participants were debriefed, solicited for comments, and paid.

Subjects

Sixty-four subjects (48 males and 16 females) ranging in age between 18 and 37 years (mean = 22.2) participated in this study. All were volunteers from the VPI & SU community and were compensated ($5/hr) for their time. On a pre-experiment questionnaire, most subjects categorized themselves as experienced PC users. However, subjects ranged in experience from basic software users to highly experienced programmers; all were familiar with the operation of mouse-based interfaces. While no specific sex differences were expected, male and female subjects were equally assigned to conditions. Data collection required 45-75 min. per subject.

Dependent Variables

This study included both performance and preference measures. Performance measures were task completion times, errors, deviations from optimum paths, time per node, and uncertainty measures. User subjective ratings and comments comprised the preference measures. Each is explained below in detail.

Task completion time (solution time). Task completion time (TCT) is the time difference between the start and end of a task. Beginning of the task was defined as the time at which the subjects entered the database, while the end was marked by the time subjects accessed the online answer sheet. All subjects read the IR task completely before entering the database; thus, TCT reflects solely the solution times with no contamination from time required to read and understand the task.

Number of erroneous responses. This measure is the number of incorrect responses to tasks. These errors were classified in a binary fashion: any answer but the correct answer is considered an error.

Deviations from the optimum path. For each task, the experimenter determined nodes considered relevant for the question. Levels of the independent variable of number of required links determined this optimum path for each task. An optimum path was defined as the shortest and most effective route through the network for obtaining the answer to each task. Subjects were anticipated to deviate, in varying degrees, from this defined path. The magnitude of these deviations was measured as:

$$\frac{\text{\# of total nodes accessed}}{\text{\# of relevant nodes accessed}}. \quad (7)$$

A low deviation ratio (the minimum value is 1.0) indicates few irrelevant nodes were accessed, while a high ratio indicates the opposite. It was presumed that a low average ratio is more desirable since it indicates efficiency and ease of locating the information.

Uncertainty measures. Frick and Miller (1951) and Miller and Frick (1949) successfully used an information theoretical approach in characterizing the predictability of response patterns of rats in a maze environment. In particular, Frick and Miller were able to evaluate contingent uncertainties in rats' behavior given the occurrence of a specific event, such as having made a trip to the food tray.

A similar approach was adopted in evaluating the patterns of link traversals in the current study. Principles of information theory were used to describe the uncertainty associated with the path used in locating the answer to a given IR task. Assume subjects start at node "A" and must locate the answer in node "Z" (for a single specific task). Monitoring subjects' movements through the network yielded a certain number of paths used to reach node "Z." Table 3 illustrates a typical sample for one task. Assuming 16 subjects per cell, the probability associated with each path, p_i, was:

TABLE 3. Sample of p(i) Values for Completing a Task. Each travelled path produces a specific probability, which is included in the uncertainty calculations.

PATH (NODE-NODE)	NUMBER OF SUBJECTS USING THIS PATH, *i*	p(i)
A - B - D - H - Z	2	0.125
A - C - I - H - Z	5	0.3125
A - B - J - K - B - D - H - Z	7	0.4375
A - C - E - F - E - C - I - H - Z	2	0.125
	16	1.0000

$$p_i = \frac{\text{\# of subjects using path i}}{\text{total \# of subjects in cell}} . \tag{8}$$

The previous example in Table 3 yielded 4 different paths with probabilities of 0.125, 0.3125, 0.4375, and 0.125, respectively. The total derived uncertainty for answering this question (getting from "A" to "Z") in this particular system configuration can be calculated in terms of bits of information:

$$\mathbf{U} = -\sum_{i=1}^{n} p_i \log_2 p_i, \tag{9}$$

where U is the total uncertainty, in bits, associated with the path taken from "A" to "Z," n is the number of different paths, and p_i is the probability of a given path. The assumption is that variability of navigation (used paths) should be reduced for linking configurations that facilitate finding information. These values can be compared statistically using standard Chi-squared tests. Attneave (1959) describes the relationship between T, the estimate of difference in information between two events (T = Uncertainty$_1$ - Uncertainty$_2$), and χ^2 as:

$$\chi^2 \approx 2 (\ln 2)\, n\, T, \tag{10}$$

where n is the sample size (or number of observations). Multiple pair-wise comparisons were made for the independent variables and their interactions. The following formulation was used to control experiment-wise alpha error:

$$\alpha_{EW} = 1 - (1 - \alpha_p)^c , \tag{11}$$

where α_{EW} is the experiment-wise alpha error (at 0.05), *c* is the number of comparisons, and α_p is the alpha level for each comparison.

At first glance uncertainty measures and deviation ratios may appear quite similar; however, while both are measures of navigational difficulty, they actually measure different aspects of such difficulty. A deviation ratio provides a measure of the degree of deviation from an author-defined optimum path. While this path for reaching an answer may, indeed, be the quickest route to task completion, it may not be the most obvious or most traveled path. Hence, the uncertainty measure will provide an understanding of the number of different paths taken in answering a question and an information theoretical description of the proportion of subjects using each path.

Time per node. Time per node is the average time spent at each node during task performance. This measure is defined by:

$$T = \frac{\text{Task time for task 1}}{\text{\# of nodes visited}}, \tag{12}$$

where T is the time per node for task 1.

Subjective measures. The Questionnaire for User Interface Satisfaction (QUIS) version 5.0 developed by the University of Maryland (Chin, Diehl, and Norman, 1988) was used as the basic questionnaire. QUIS developers report high reliability factors (Cronbach's alpha = 0.94) and the HCI Laboratory at the University of Maryland has reported successful uses of QUIS (Shneiderman, 1987). Subjects provided ratings of 1 to 9 (1 being low or negative and 9 being high or positive) for the different linking structures on dimensions included in

QUIS (Appendix V). These ratings were on an absolute basis since each subject was exposed to only one level of linking structure. The QUIS was slightly modified to exclude a completely irrelevant question and add five questions believed to be important to the specific aims of this research. The original QUIS was supplemented by questions 26, 27, 28, 29, and 30 as presented in the appendix. Obtained ratings were treated as interval scale data and a MANOVA was used in their analysis. A MANOVA is appropriate because one may safely assume a high degree of correlation among the questionnaire items as suggested by the high Cornbach's alpha value.

RESULTS

This chapter discusses the statistical analyses and results for data collected in Experiments A and B. These data were collected simultaneously; however, a priori distinctions between the experiments allowed for separate analyses of all objective measures. No differentiation is made between Experiments A and B for the subjective questionnaire since subjects completed the questionnaire subsequent to performing all tasks.

Experiment A

Experiment A contains the results of IR tasks requiring equivalent numbers of links on the network and hierarchical systems. From a task performance view, an optimum information retrieval system will yield low task times and errors, as well as low deviation ratios, times per screen, and uncertainty values.

Task completion time (TCT). A three-way ANOVA tested the effects of linking structure, number of required links, task type, and their interactions on TCT. Table 4 presents the ANOVA summary table. The Greenhouse-Geisser ε correction factor for violation of sphericity was calculated for all within-subjects variables and concomitant interactions; corrected alpha levels appear in the table for these sources of variance (Greenhouse-Geisser ε is not required for within-subjects variables with two levels). Refer to Greenhouse and Geisser (1959) and Huynh and Feldt (1970) for further detail on ε.

TABLE 4. ANOVA Summary Table for TCT (Experiment A). The included *p*-values have been corrected by the Greenhouse-Geisser (1959) ε where necessary.

SOURCE OF VARIANCE	df	SS	MS	F	p
Between-Subjects					
System (S)	3	157243.966	52414.655	4.59	0.0059
Subjects/System (S/S)	60	685243.989	11420.733		
Within-Subjects					
Links (L)	2	1084607.453	542303.727	152.25	0.0001
L x S	6	66460.464	11076.744	3.11	0.0118
L x S/S	120	427426.416	3561.887		
Task Type (T)	1	3756.253	3756.253	1.07	0.3061
T x S	3	35742.424	11914.141	3.38	0.0239
T x S/S	60	211533.823	3525.564		
L x T	2	368657.224	184328.612	53.19	0.0001
L x T x S	6	10785.612	1797.602	0.52	0.7582
L x T x S/S	120	415888.458	3465.737		
Total	383	3467346.082			

Greenhouse-Geisser Epsilon value for L and L x S = 0.8346

Greenhouse-Geisser Epsilon value for L x T and L x T x S = 0.8173

The system and link main effects yielded significant *F*-values, as did the System x Links, System x Task Type, and Links x Task Type interactions. Figures 26 and 27 present the significant main effect results. Newman-Keuls post-hoc tests (overall $\alpha = 0.05$) failed to show a significant difference between the hierarchical and combination systems (Table 5). However, these two systems yielded significantly shorter TCTs than the network and linear systems, which were found to not differ significantly from one another.

For the links main effect, a Newman-Keuls post-hoc test found 2-link task times reliably shorter than 4-link task times, both of which were significantly shorter than 6-link task times. This links main effect is intuitively satisfying and confirms the notion that tasks become more challenging (measured by TCT) as the number of required links increases. The lack of a task type main effect illustrates the equality among the task types in difficulty, indicating TCT differences among the tasks can be reliably attributed to differences in system structure and required links. Figure 28 illustrates the proximity in mean TCT and standard error for levels of task type.

Figure 29 presents the System x Links interaction. A Newman-Keuls post-hoc test (overall $\alpha = 0.05$) found no differences among the systems for 2-link tasks (Table 6). For 4-link tasks the linear system showed significantly longer task completion times than all other systems, which failed to differ significantly. Finally, for 6-link tasks the linear and network systems required significantly more time than the hierarchical or combination systems, with no reliable differences uncovered between linear and network nor between hierarchical and combination. It is interesting to note that TCTs for 6-link tasks on the

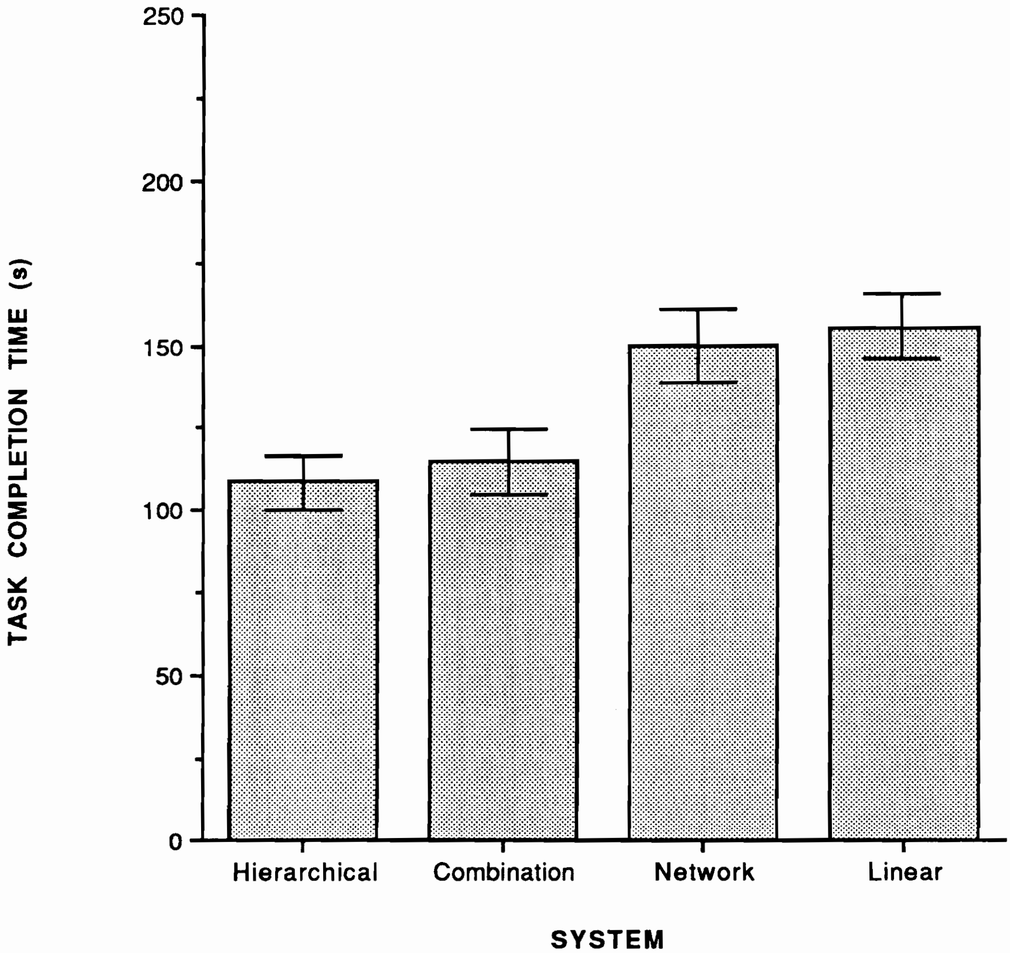


Figure 26. System configuration main effect on task completion time (Experiment A). Standard error bars are included for each system.

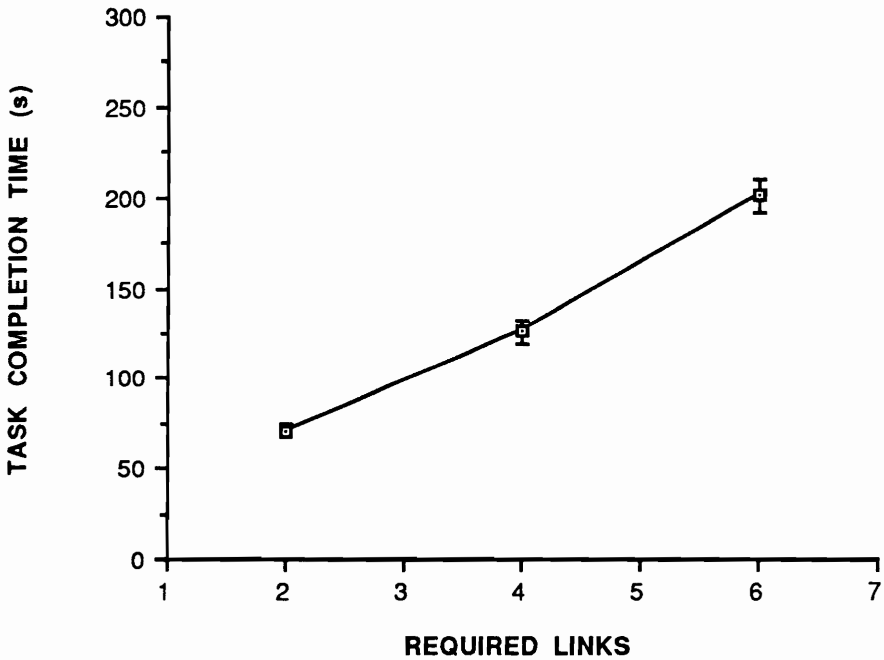


Figure 27. Effect of number of required links on task completion time. Overall regression equation for task completion time as a function of number of required links is *TCT* = 2.8287 + 32.412 *Links*, R^2 = 0.992.

TABLE 5. Newman-Keuls Post-Hoc Test for TCT System Main Effect ($p < 0.05$) in Experiment A. TCTs with identical letters are not significantly different.

SYSTEM	TCT (s)	
Hierarchical	109.21	A
Combination	115.78	A
Network	149.35	B
Linear	155.56	B

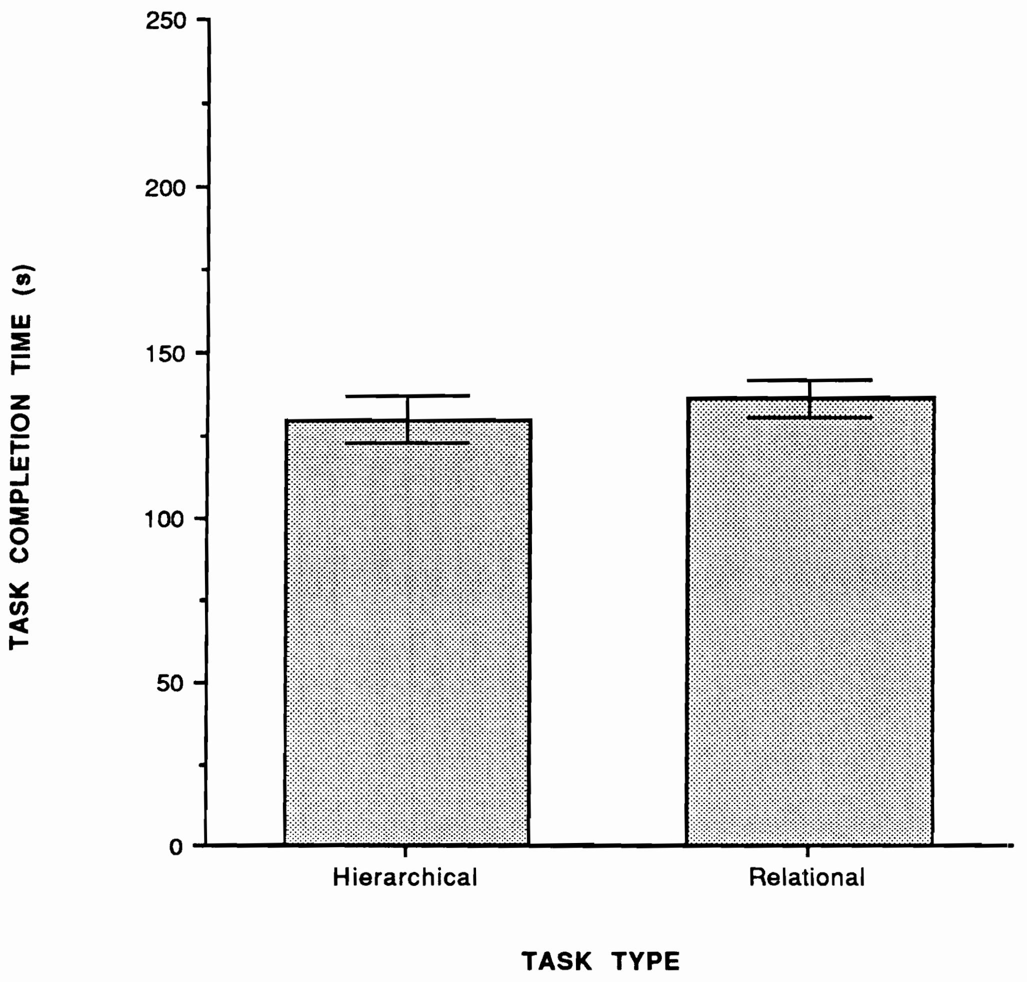


Figure 28. Effect of task type on task completion time (no significant difference exists).

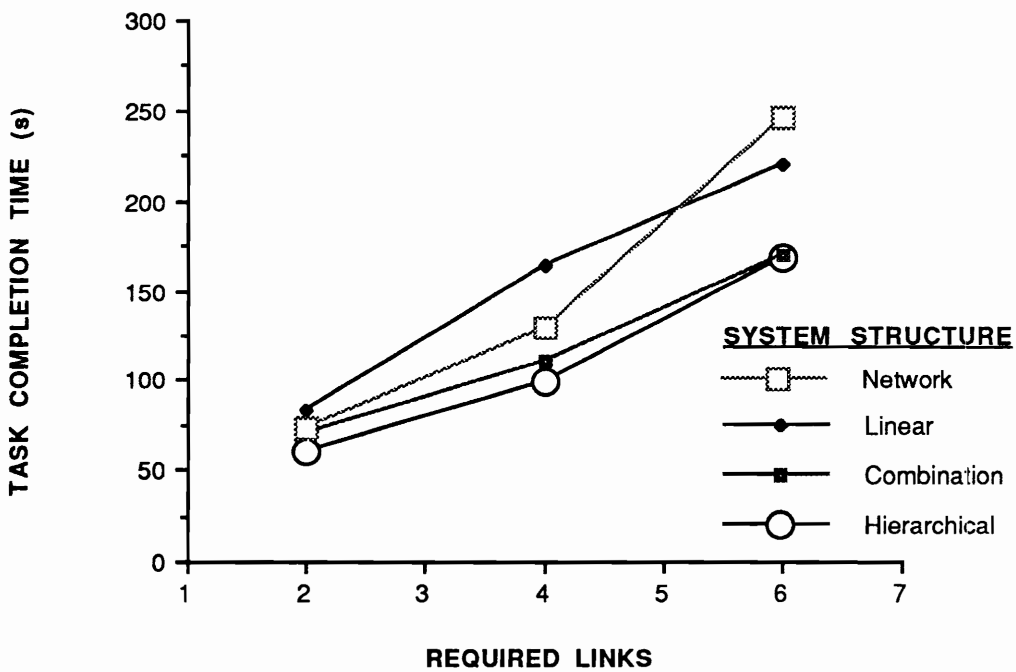


Figure 29. System x Links interaction for task completion time.

TABLE 6. Newman-Keuls Post Hoc Test for TCT System x Required Links Interaction ($p < 0.05$) Experiment A. TCTs with identical letters are not significantly different.

SYSTEM	REQUIRED LINKS	TCT (s)		
Hierarchical	2-link	60.88	A	
Combination	2-link	70.97	A	
Network	2-link	73.88	A	
Linear	2-link	83.44	A	
Hierarchical	4-link	99.50	A	B
Combination	4-link	110.34	A	B
Network	4-link	129.31	C	B
Linear	4-link	163.60	C	
Hierarchical	6-link	167.30	C	
Combination	6-link	169.10	C	
Network	6-link	219.60	D	
Linear	6-link	245.00	D	

hierarchical and combinations systems were as low as (no reliable difference) TCTs for 4-link tasks on the linear and network systems.

Regression equations for the TCT means were computed for each system as a function of required links. Figure 30 presents the best-fit lines for each system. The following *t*-test formulation presented by Montgomery (1984) for hypothesis testing in linear regression was used to test for differences among the β-weights for the different equations:

$$t = \frac{\beta - \beta_j}{\sqrt{MS_E / S_{xx}}} . \tag{13}$$

In this equation β is the obtained β-weight for a line, β_j is the value β is tested against, MS_E is the residual mean square, and S_{xx} is the corrected sum of squares of the x variable (i.e., links). The experiment-wise α error (0.05) was adjusted for six paired comparisons using equation 11; the corresponding tabled *t*-value was used to test for significance of differences. The network system exhibited higher β-weight values than the hierarchical ($t = 2.729$, $p < 0.0085$) and combination systems. No other differences were uncovered.

The System x Task Type interaction is depicted in Figure 31. As seen in Table 7, for relational tasks a Newman-Keuls test (overall $\alpha = 0.05$) failed to discover differences between the combination and hierarchical systems, and between the network and hierarchical systems. However, the linear system yielded significantly longer TCTs than all others, while the combination system produced significantly shorter TCT than the network configuration. Of particular

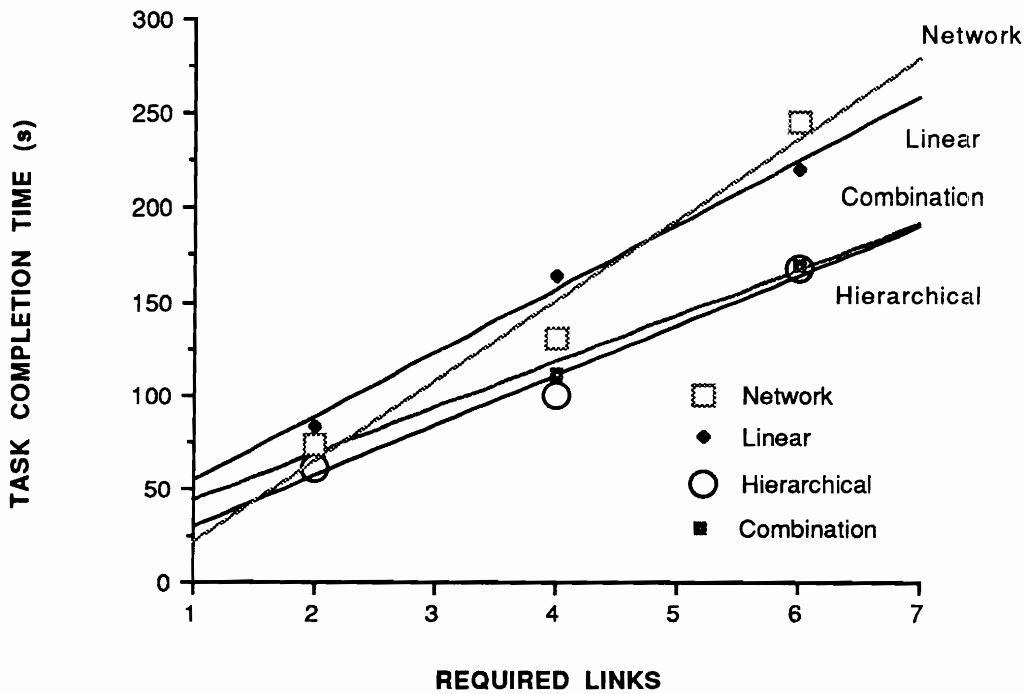


Figure 30. Regression lines fit to the the mean required link TCTs for each system. The linear regression equations for each system are:

Linear -- $TCT_L = 19.375 + 34.05\ \mathit{Links},\ r^2 = 0.99$

Hierarchical -- $TCT_H = 2.80 + 26.60\mathit{Links},\ r^2 = 0.975$

Network -- $TCT_N = -21.896 + 42.813\mathit{Links},\ r^2 = 0.961$

Combination -- $TCT_C = 11.031 + 26.188\mathit{Links},\ r^2 = 0.987$

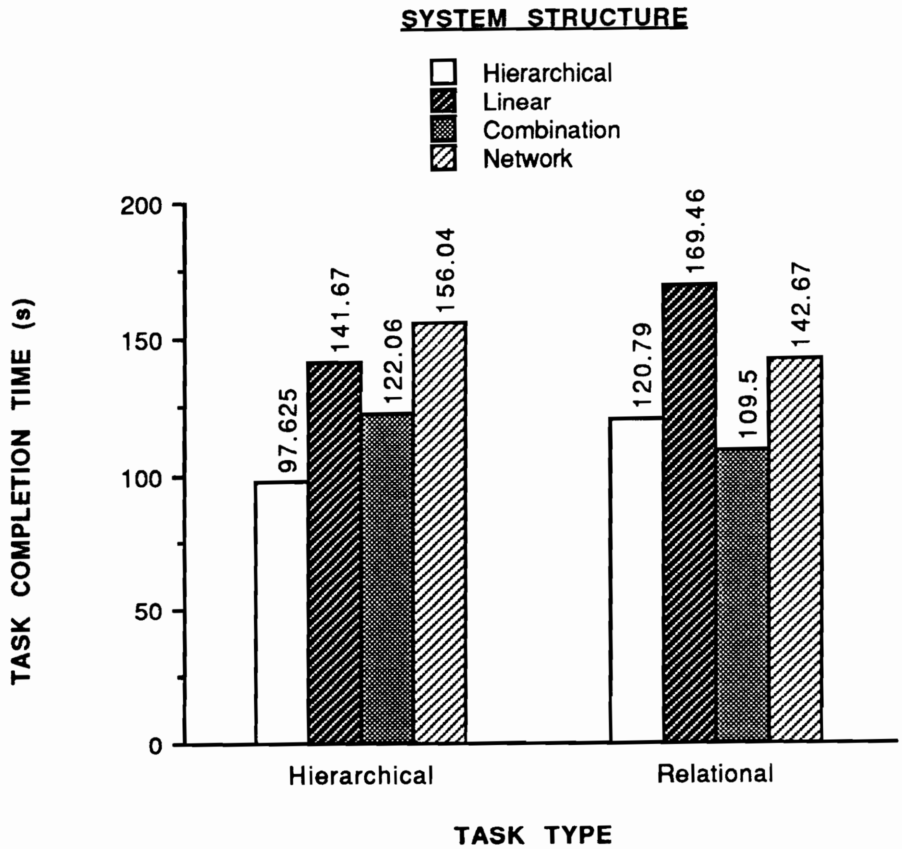


Figure 31. The System x Task Type interaction effect for task completion time.

TABLE 7. Newman-Keuls Post Hoc Test for TCT System x Task Type Interaction ($p < 0.05$) in Experiment A. TCTs with identical letters are not significantly different.

SYSTEM	TASK TYPE	TCT (s)			
Hierarchical	Hierarchical	97.63	A		
Combination	Relational	109.50	A		
Hierarchical	Relational	120.79	A	B	
Combination	Hierarchical	122.06	A	B	
Linear	Hierarchical	141.67		B	C
Network	Relational	142.67		B	C
Network	Hierarchical	156.04			C
Linear	Relational	169.40			C

interest is the absence of a difference between the network and hierarchical systems. Recall that expert judges determined relational tasks to be particularly well suited to the network configuration; this result indicates that the hierarchical configuration can perform as well as the network for such relational tasks. The combination structure yielded the best results for relational tasks. For hierarchical tasks no differences were found between the network and linear systems nor between the hierarchical and combination systems; however, the former pair produced significantly longer TCTs than the latter pair. Hence, the network configuration yields detrimental effects for hierarchical tasks as compared to the hierarchical and combination systems.

The Task Type X Link interaction also produced a significant *F*-value. Figure 32 illustrates the obtained mean TCTs and Appendix VI presents the Newman-Keuls post-hoc test. Note that the plotted TCTs are those associated with individual tasks; all tasks are defined by one level of task type and one level of required links. Post-hoc tests (Newman-Keuls overall $\alpha = 0.05$) failed to detect a difference among 2-link hierarchical, 2-link relational, and 4-link hierarchical tasks, all of which exhibited significantly shorter TCTs than the remaining tasks. No difference was found between the 4-link and 6-link relational tasks, but both had significantly shorter TCTs than the 6-link hierarchical task. Collapsed across task type, the data exhibit logical trends as found in the required links main effect. The lack of a task type main effect also illustrates that these data show no TCT differences when collapsed across required links, again an intuitively satisfying and desired result. These trends are consistent across systems as determined by the absence of a third-order

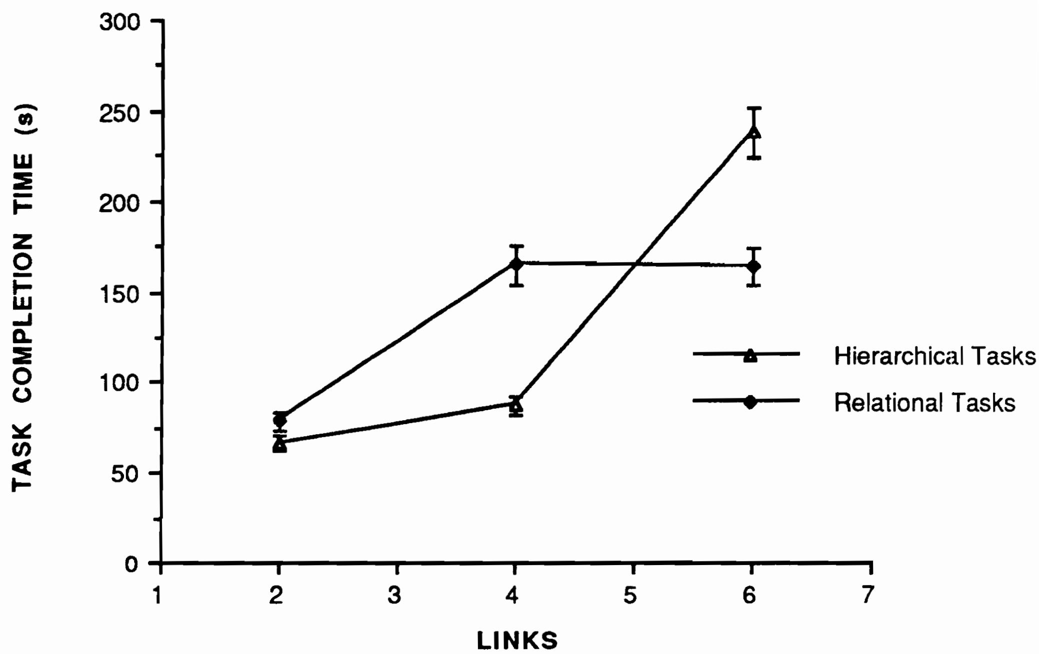


Figure 32. The Task Type x Link interaction effect on task completion time.

interaction; hence, the obtained Task Type X Link interaction must be attributed to inherent differences among the tasks.

Errors. Table 8 provides a matrix of the frequency of errors. A total of six IR tasks were performed on each of four different structures 16 times yielding a total of 384 tasks. However, only 18 errors occurred resulting in a 4.6 % error rate. The frequency of errors for the systems is quite low and no statistical analyses were conducted. It is premature to claim errors were insensitive to the independent variables; rather, this author believes subjects were highly motivated to retrieve the correct answer at the expense of increased TCT. Three observations lead to this statement: (1) subjects claimed to enjoy using the systems, (2) most all subjects found the information domain interesting and were determined to fulfill experimental requirements by providing the correct answer (many indicated they would have better enjoyed high school geography and history had they been able to use such versatile and attractive systems), and (3) due to the absence of specific time limits, subjects felt compelled to uncover the correct responses. Time limitations may have yielded considerably different error results.

Uncertainty. Attneave's (1959) bits to χ^2 transformation (formula 10) was used to investigate the differences in uncertainty. Figure 33 illustrates the uncertainty associated with each of the three systems of interest. A test of the differences in uncertainty among the systems found no significant difference between the hierarchical and combination systems, but both yielded significantly less uncertainty than the network system ($p < 0.0169$ for three paired comparisons). Additionally, 2-link tasks had lower associated

TABLE 8. Errors for each System in Experiment A.

	Hierarchical Tasks			Relational Tasks		
	2-links	4-links	6-links	2-links	4-links	6-links
Linear	0	1	3	0	0	0
Hierarchical	0	0	1	0	2	0
Network	0	1	4	0	0	1
Combination	3	0	0	0	1	1

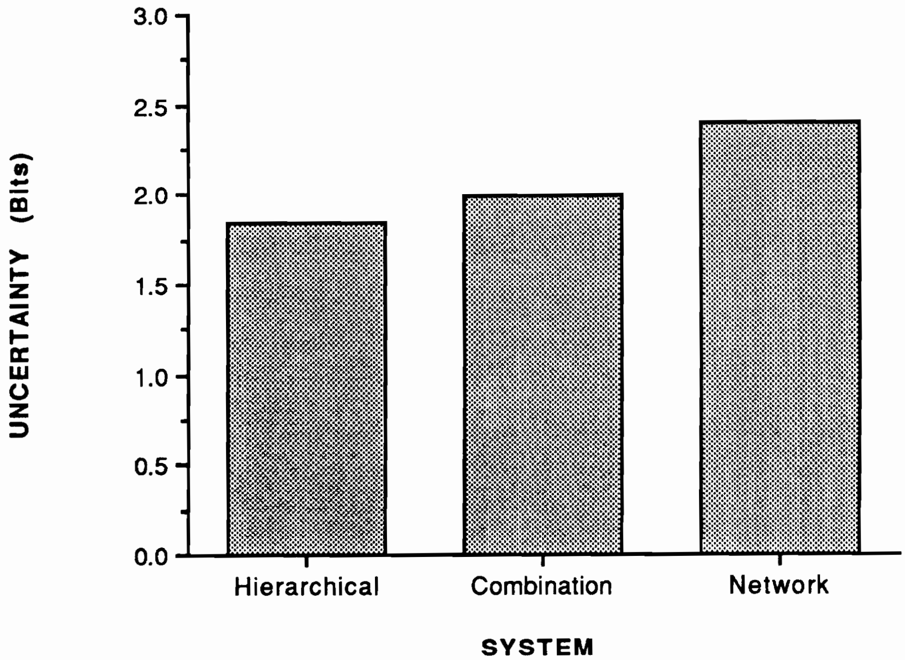


Figure 33. Effect of system configuration on uncertainty.

uncertainty than 4-link tasks, both of which exhibited lower uncertainty than 6-link tasks (Figure 34). Task type failed to produce a significant difference in uncertainty.

Figure 35 presents the System x Links interaction. No differences exist among the systems for 2- and 4- link tasks; however, the network configuration yields significantly higher uncertainty than the others for 6-link tasks ($p < 0.005$ for nine paired comparisons (overall $\alpha = 0.05$). The hierarchical and combination systems exhibit no differences at any level of required links. The level of similarity between TCTs and uncertainty for this interaction is evident. Figure 36 illustrates mean task completion times plotted against derived uncertainties for each task (across the 3 systems). A strong correlation exists between TCT and uncertainty, $r = 0.858$. Uncertainty was not a controlled variable; hence, no conclusions of causality are appropriate. However, the strong correlation between these measures allows one to confidently predict navigational difficulties given knowledge of TCT, and vice versa. This relationship is robust for all system structures. An analysis of the Task Type x Links interaction failed to produce differences between links at various levels of task type ($p < 0.0083$ for six paired comparisons, overall $\alpha = 0.05$).

Deviations from the optimum path. A three-way ANOVA was used for analysis of the deviation data. Table 9 includes the resulting ANOVA summary table. As with TCTs, the Greenhouse-Geisser ε correction factor for violation of sphericity was calculated for all within-subjects variables and associated interactions; corrected alpha levels appear in the table for these sources of variance.

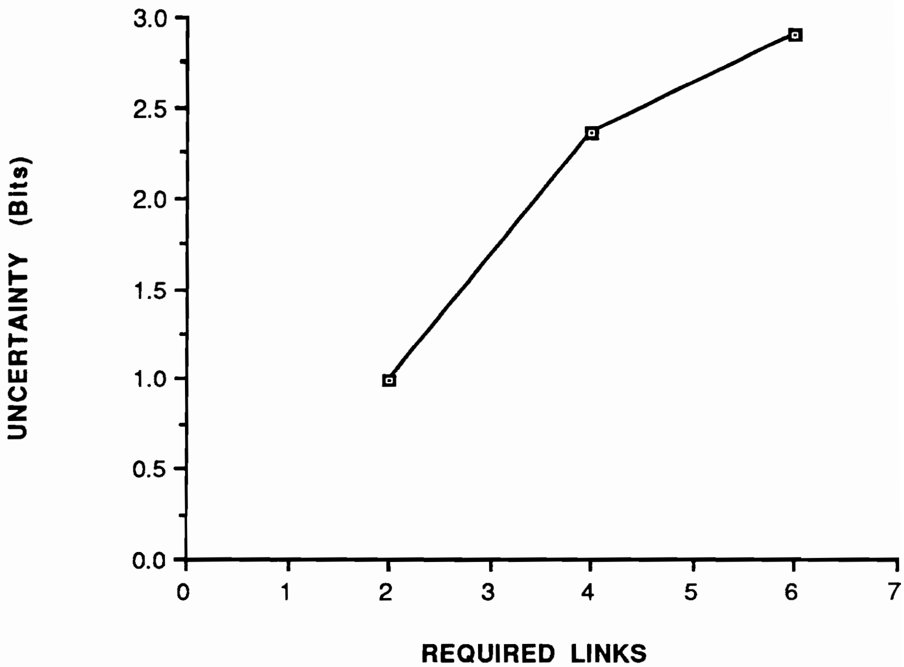


Figure 34. Effect of number of required links on uncertainty.

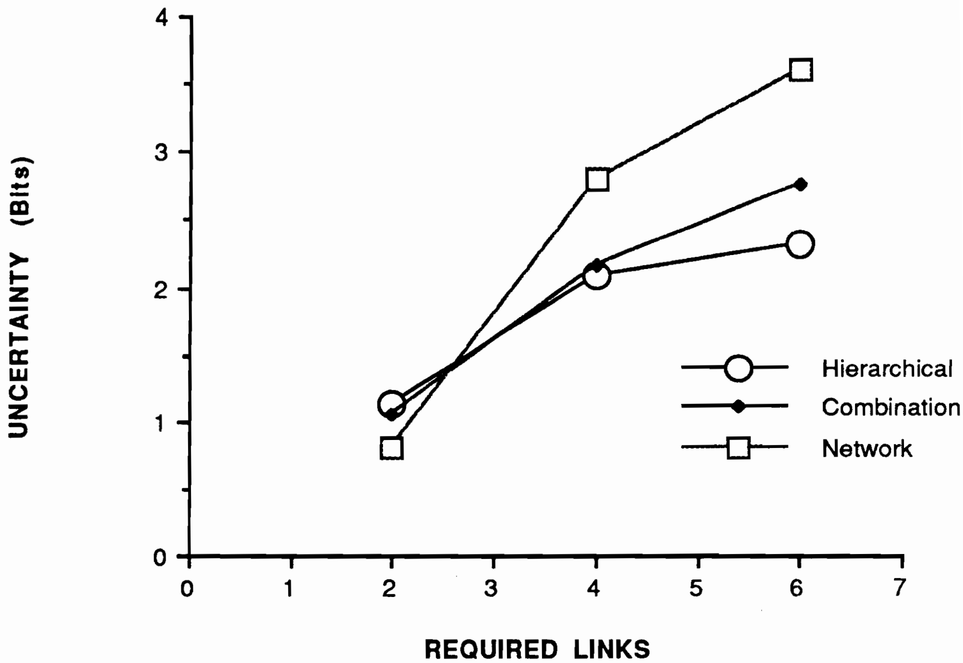


Figure 35. The System x Link interaction for uncertainty.

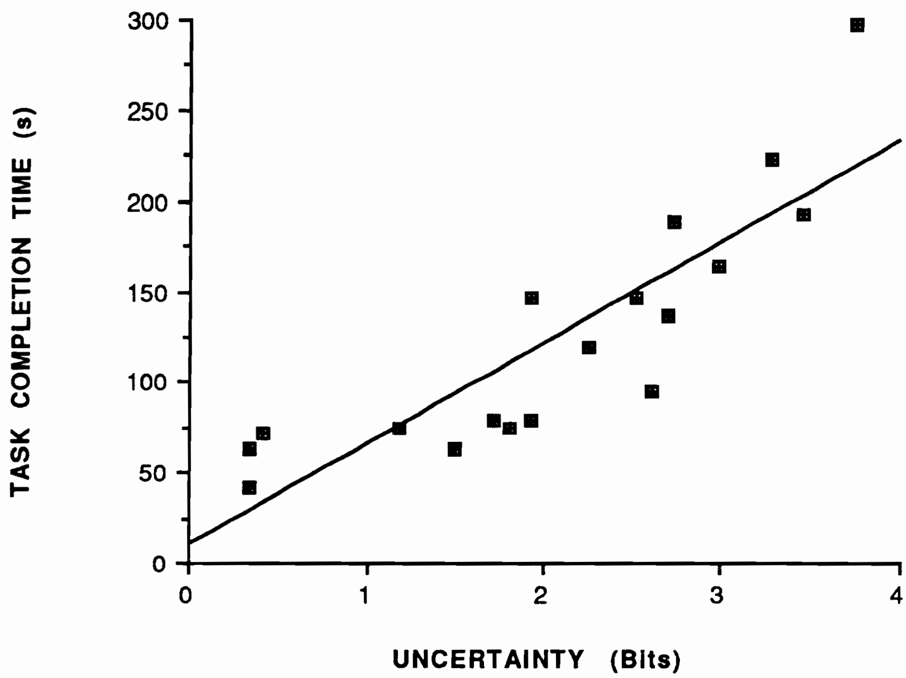


Figure 36. Task completion times plotted against associated uncertainty for each task performed on the systems (*r* = 0.858).

TABLE 9. ANOVA Summary Table for Deviations (Experiment A). The included *p*-values have been corrected by the Greenhouse-Geisser (1959) ε where necessary.

SOURCE OF VARIANCE	df	SS	MS	F	p
Between-Subjects					
System (S)	2	16.862	8.431	4.63	0.0149
Subjects/System (S/S)	45	81.969	1.822		
Within-Subjects					
Links (L)	2	0.363	0.182	0.16	0.8042
L x S	4	8.796	2.199	1.99	0.1183
L x S/S	90	99.505	1.106		
Task Type (T)	1	2.971	2.971	2.87	0.0971
T x S	2	12.192	6.096	5.89	0.0053
T x S/S	45	46.554	1.035		
L x T	2	16.228	8.114	6.24	0.0067
L x T x S	4	7.685	1.921	1.48	0.2278
L x T x S/S	90	116.937	1.299		
Total	287	410.062			

Greenhouse-Geisser Epsilon value for L and L x S = 0.8123

Greenhouse-Geisser Epsilon value for L x T and L x T x S = 0.7563

Task Type and Required Links failed to produce significant main effects, but the System main effect is significant. Figure 37 presents the system effect for deviation. A Newman-Keuls post-hoc test, presented in Table 10, found the hierarchical system to have significantly higher deviation than the network and combination systems, which are not significantly different. At first glance this finding seems somewhat inconsistent with the results for TCT and uncertainty: while the hierarchical configuration produces lower TCT and uncertainty, it also exhibits significantly higher deviation. The System x Task Type interaction (Figure 38) further explains this seemingly contradictory result. No differences exist among the mean system deviations for hierarchical tasks at an overall α of 0.05 (Table 11). These mean deviations are not reliably different from those of the network and combination systems for relational tasks; however, the hierarchical system yields significantly higher deviation for relational tasks. Moreover, this condition seems consistent across number of required links as indicated by the absence of a third-order interaction. Detailed explanations of this finding will be included in the Discussion and Conclusions chapter.

Figure 39 presents the Link x Task Type interaction. Only the 2-link relational deviation mean is significantly different from the 2-link and 4-link hierarchical and the 6-link relational tasks (overall $\alpha = 0.05$). The remaining five means are not reliably different from one another.

Time per screen (Node). A three-way ANOVA was used to analyze time per screen data; Table 12 presents the summary table. Post-hoc analysis of the system configuration main effect (Table 13) found the network system to have

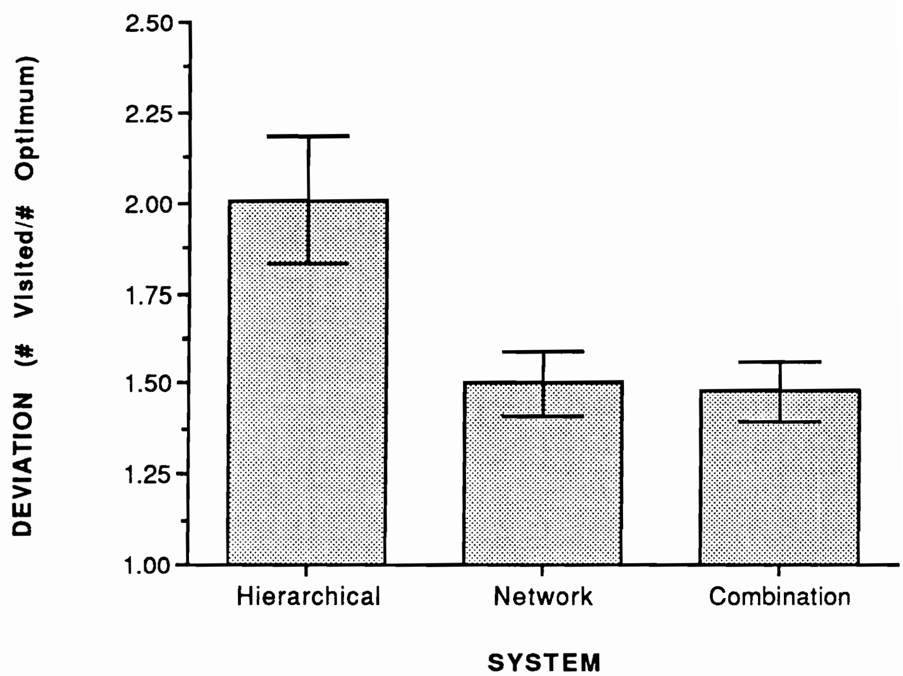


Figure 37. Deviation system main effect.

TABLE 10. Newman-Keuls Post Hoc Test for Deviaton System Main Effect ($p < 0.05$) in Experiment A. Deviations with identical letters are not significantly different.

SYSTEM	DEVIATION	
Combination	1.484	A
Network	1.503	A
Hierarchical	2.006	B

NOTE: Deviation is measured as number of visited nodes / number of required links.

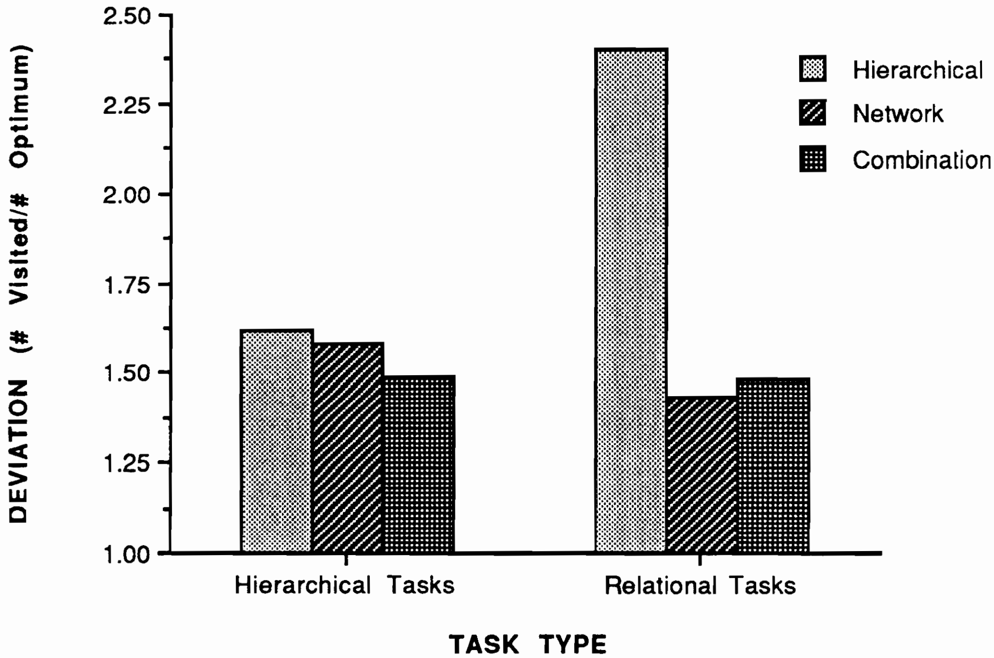


Figure 38. System x Task Type interaction for deviation measures.

TABLE 11. Newman-Keuls Post Hoc Test for Deviation System x Task Type Interaction ($p < 0.05$) in Experiment A. Deviations with identical letters are not significantly different.

SYSTEM	TASK TYPE	DEVIATION	
Network	Relational	1.425	A
Combination	Relational	1.476	A
Combination	Hierarchical	1.491	A
Network	Hierarchical	1.580	A
Hierarchical	Hierarchical	1.616	A
Hierarchical	Relational	2.396	B

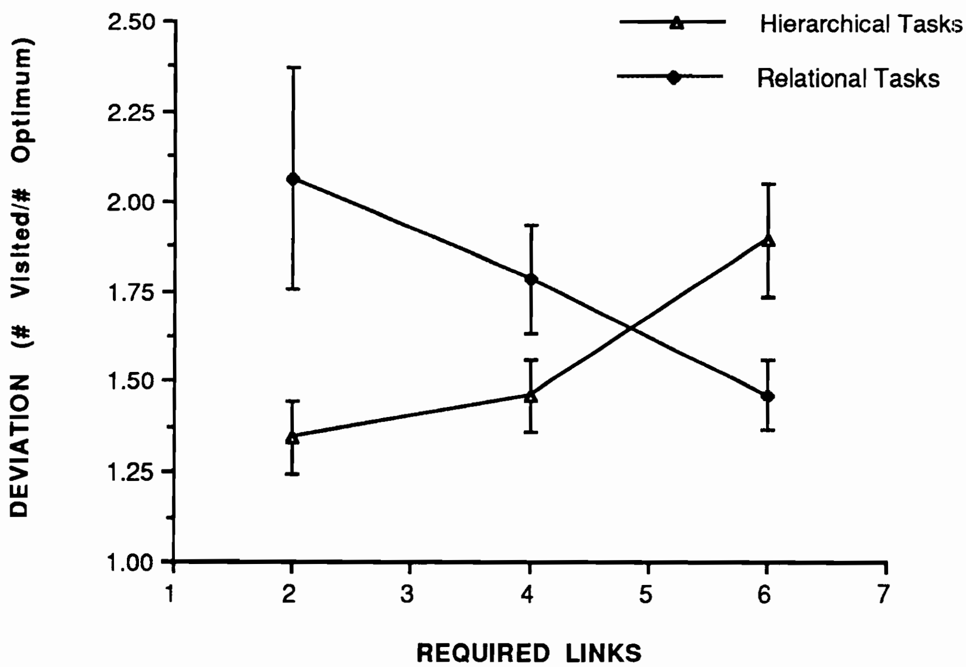


Figure 39. The Task Type x Link interaction for deviation.

TABLE 12. ANOVA Summary Table for Time Per Screen (Experiment A). The included *p*-values have been corrected by the Greenhouse-Geisser (1959) ε where necessary.

SOURCE OF VARIANCE	df	SS	MS	F	p
Between-Subjects					
System (S)	2	3420.656	1710.328	6.56	0.0032
Subjects/System (S/S)	45	11734.7	260.771		
Within-Subjects					
Links (L)	2	1302.831	651.416	11.23	0.0001
L x S	4	431.708	107.927	1.86	0.1323
L x S/S	90	5219.837	57.998		
Task Type (T)	1	3.819	3.819	0.07	0.7977
T x S	2	5.762	2.881	0.05	0.9511
T x S/S	45	2583.813	57.418		
L x T	2	2585.349	1292.675	26.60	0.0001
L x T x S	4	485.053	121.263	2.50	0.0599
L x T x S/S	90	4373.898	48.599		
Total	287	32147.42			

Greenhouse-Geisser Epsilon value for L and L x S = 0.8967

Greenhouse-Geisser Epsilon value for L x T and L x T x S = 0.8375

TABLE 13. Newman-Keuls Post Hoc Test for Time per Node System Main Effect ($p < 0.05$) in Experiment A. Times with identical letters are not significantly different.

SYSTEM	TIME/NODE (s)		
Hierarchical	16.652	A	
Combination	21.297	A	B
Network	25.074		B

significantly longer mean time per screen than the non-differing hierarchical and combination systems (overall $\alpha = 0.05$). Figure 40 Illustrates these findings. The required links main effect also produced a significant *F*-value. Two-link tasks exhibited significantly longer times per screen than either 4- or 6-link tasks, which failed to yield a significant difference (Figure 41).

The sole higher-order interaction found to be significant is the Task Type x Links interaction. This result is presented in Figure 42. Means depicted in the graph are the times per screen associated with each IR task. No differences were found among 2-link hierarchical, 2-link relational, 4-link relational, and 6-link hierarchical tasks; all of these tasks exhibit significantly longer times per screen than the 6-link relational and 4-link hierarchical tasks. The 6-link relational task was found to have longer times per screen than the 4-link hierarchical.

Experiment B

Experiment B contains results of IR tasks requiring half as many links on the network system. All dependent measures are listed separately; associated analyses and results are discussed in detail.

Task completion time (TCT). A two-way ANOVA tested the effects of linking structure, number of required links, and their interaction on TCT. Table 14 presents the ANOVA summary table. The Greenhouse-Geisser ε correction factor for violation of sphericity was calculated for the required links variable and concomitant interactions; corrected alpha levels appear in the table for these sources of variance.

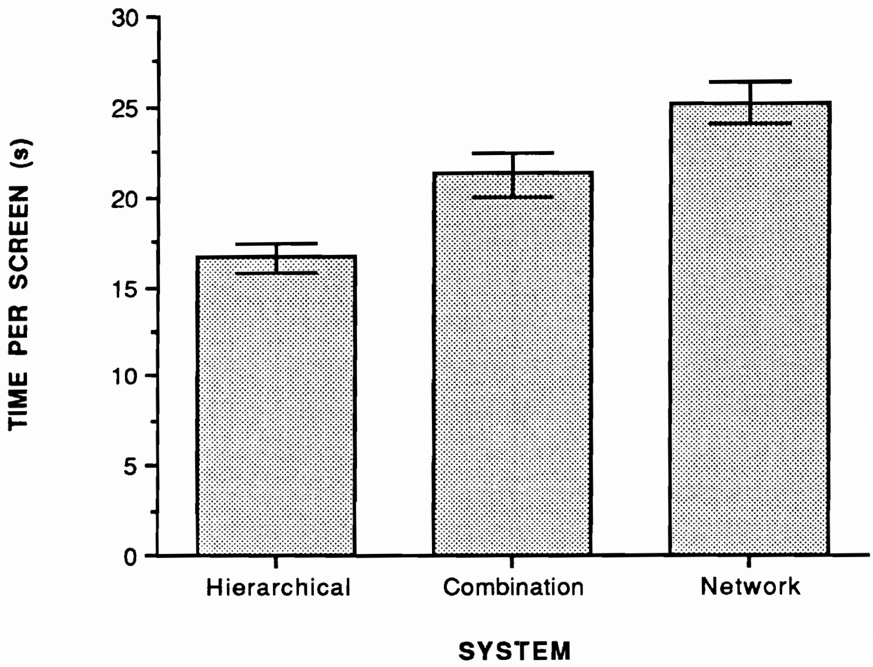


Figure 40. System effect for time per screen (node).

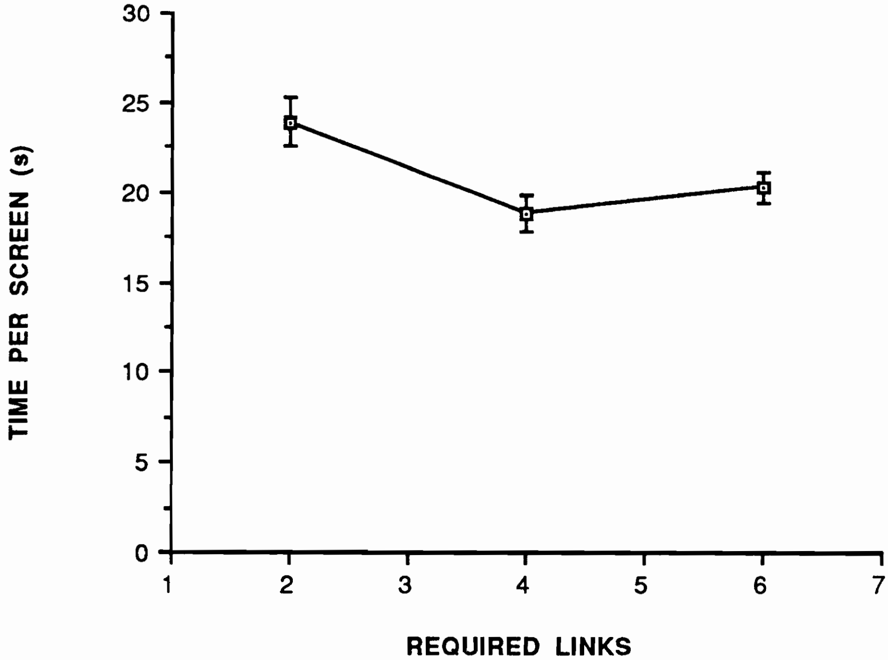


Figure 41. Required links effect for time per screen in Experiment A. No difference was uncovered between 4- and 6-links tasks, but 2-link tasks exhibited higher time per screen than the others.

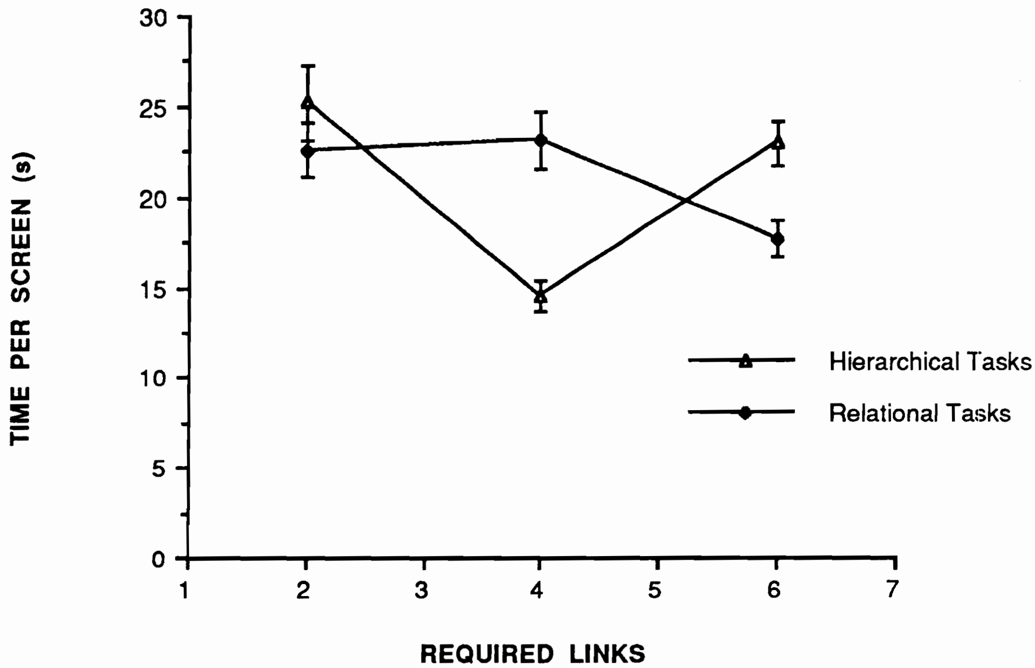


Figure 42. The Task Type x Links interaction effect for time per screen.

TABLE 14. ANOVA Summary Table for TCT (Experiment B). The included *p*-values have been corrected by the Greenhouse-Geisser (1959) ε where necessary.

SOURCE OF VARIANCE	df	SS	MS	F	p
Between-Subjects					
System (S)	3	85241.932	28413.977	6.57	0.0006
Subjects/System (S/S)	60	259508.06	4325.134		
Within-Subjects					
Links (L)	2	246553.13	123276.563	55.49	0.0001
L x S	6	50188.208	8364.701	3.77	0.0033
L x S/S	120	266582	2221.517		
Total	191	908073.33			

Greenhouse-Geisser Epsilon value for L and L x S = 0.8530

The System and Required links main effects yielded significant *F*-values, as did the System x Links interaction. Figures 43 and 44 present the main effect results. Newman-Keuls post-hoc tests (overall $\alpha = 0.05$) failed to show a significant difference between the following pairs of systems: linear - hierarchical, hierarchical - combination, and combination -network (Table 15). Thus, the network configuration yields significantly shorter TCTs than both the hierarchical and linear systems. No differences were uncovered between 2- and 4-link tasks, yet 6-link tasks produces significantly longer TCTs.

Figure 45 illustrates the System x Links interaction. A Newman-Keuls post-hoc analysis revealed that for 2-link tasks the hierarchical, network and combination systems show no difference, but the linear configuration yields significantly longer TCTs (Table 16). No differences were detected for 4-link tasks. For 6-link tasks the hierarchical and linear systems produce no difference; however, both have significantly longer TCTs than the combination system. The combination system exhibits longer mean TCT than the network system. TCT differences are absent across all levels of required links for the network system; this result is not surprising since the number of required links for this system are 1, 2, and 3 links.

Errors. Exceptionally few (total of 5) errors were committed. As explained in the results for Experiment A, subjects were highly motivated to retrieve the correct answer at the expense of increased TCT. Therefore, no analysis of errors is warranted.

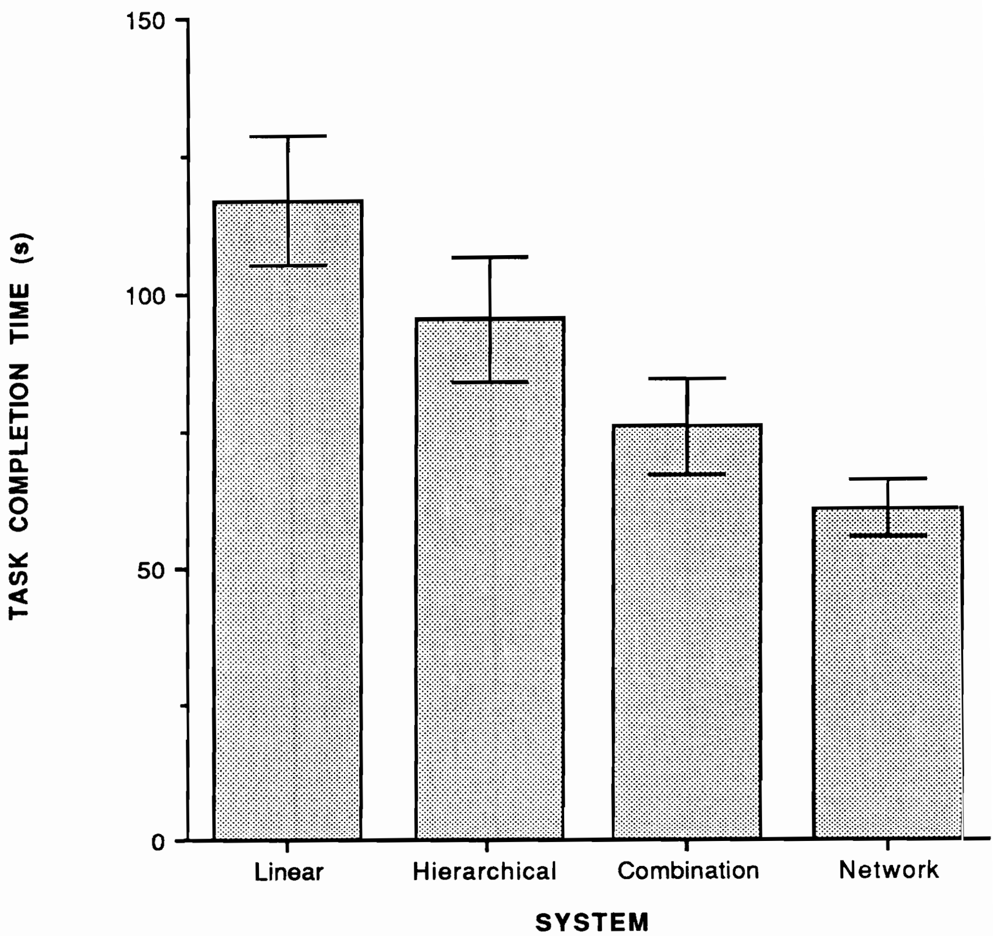


Figure 43. System effect for task completion time in Experiment B.

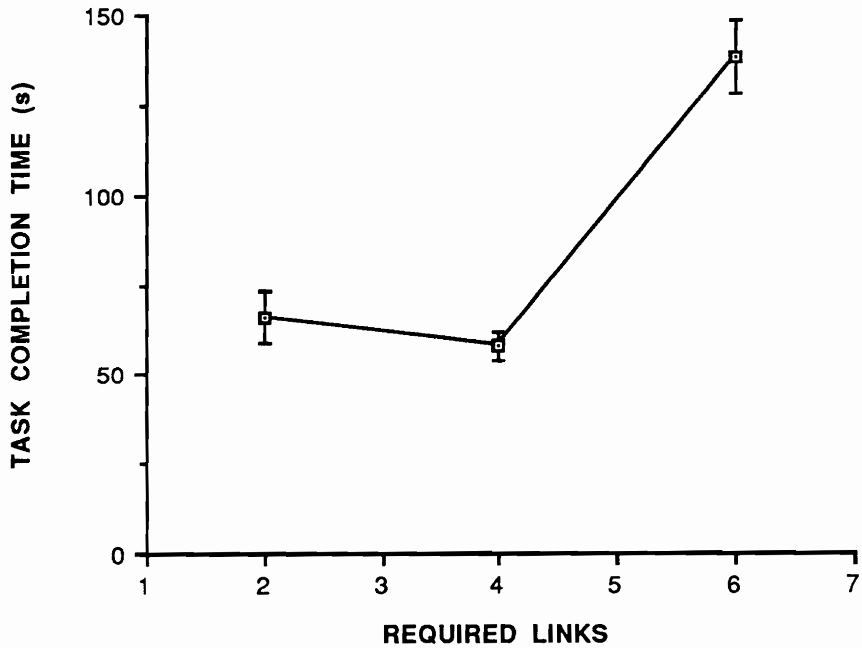


Figure 44. Required links effect for task completion time in Experiment B.

TABLE 15. Newman-Keuls Post Hoc Test for TCT System Main Effect ($p < 0.05$) in Experiment B. TCTs with identical letters are not significantly different.

SYSTEM	TCT (s)		
Network	60.65	A	
Combination	75.88	A	B
Hierarchical	95.38	C	B
Linear	116.79	C	

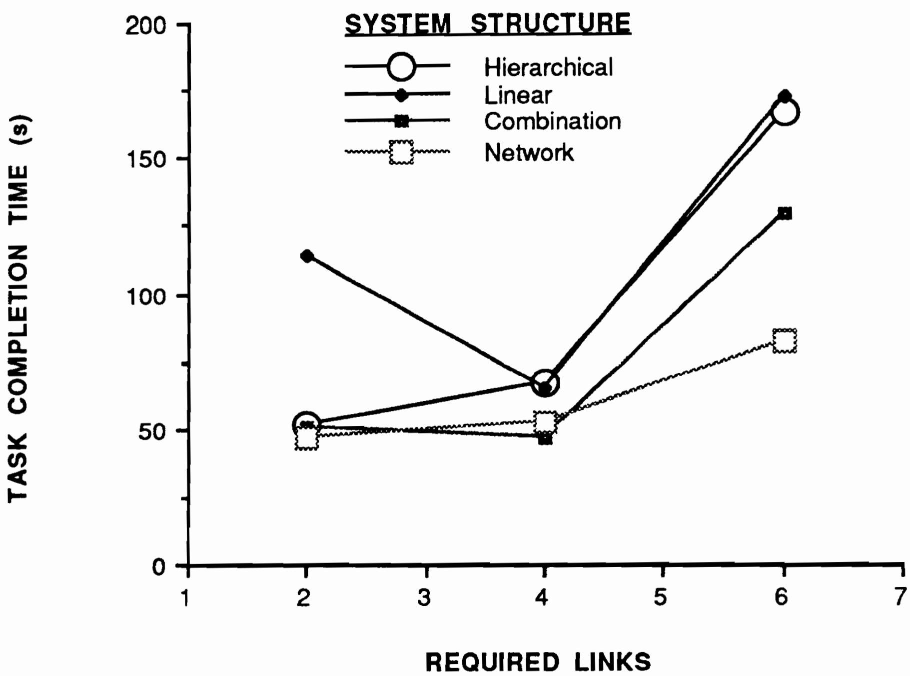


Figure 45. System x Link interaction for task completion time in Experiment B.

TABLE 16. Newman-Keuls Post Hoc Test for TCT System x Required Links Interaction ($p < 0.05$) in Experiment B. TCTs with identical letters are not significantly different.

SYSTEM	REQUIRED LINKS	TCT (s)		
Combination	4-link	46.94	A	
Network	2-link (i.e., 1-link)	47.13	A	
Combination	2-link	51.13	A	
Hierarchical	2-link	52.00	A	
Network	4-link (i.e., 2-link)	52.50	A	
Linear	4-link	64.94	A	
Hierarchical	4-link	67.44	A	
Network	6-link (i.e., 3-link)	82.31	A	B
Linear	2-link	113.44	C	B
Combination	6-link	129.56	C	
Hierarchical	6-link	166.69	D	
Linear	6-link	172.00	D	

Uncertainty. Figure 46 illustrates the uncertainty associated with each of the three systems of interest. A test of the differences in uncertainty among the systems found no significant difference between the hierarchical and combination systems, nor between the hierarchical and network systems; however, the combination system yields significantly higher uncertainty than the network configuration ($p < 0.0169$ for three paired comparisons). Additionally, 2-link tasks have lower associated uncertainty than 4-link tasks, both of which exhibit lower uncertainty than 6-link tasks (Figure 47).

Figure 48 presents the System x Links interaction. No differences exist among the systems for 4- or 6-link tasks; however, the combination configuration yields significantly higher uncertainty than the others for 2-link tasks ($p < 0.005$ for nine paired comparisons). The hierarchical and network systems exhibited no differences at any level of required links.

The correlation between TCTs and uncertainty for these tasks was investigated. Figure 49 illustrates mean task completion times plotted against derived uncertainties for each task (across the three systems). As with results from Experiment A, a strong correlation exists between TCT and uncertainty, $r = 0.848$.

Deviations from the optimum path. A two-way ANOVA was used for analysis of the deviation data. Table 17 includes the resulting ANOVA summary table. Greenhouse-Geisser ε correction factor for violation of sphericity was calculated for the within-subjects variable and its associated interactions; corrected alpha levels appear in the table for these sources of variance.

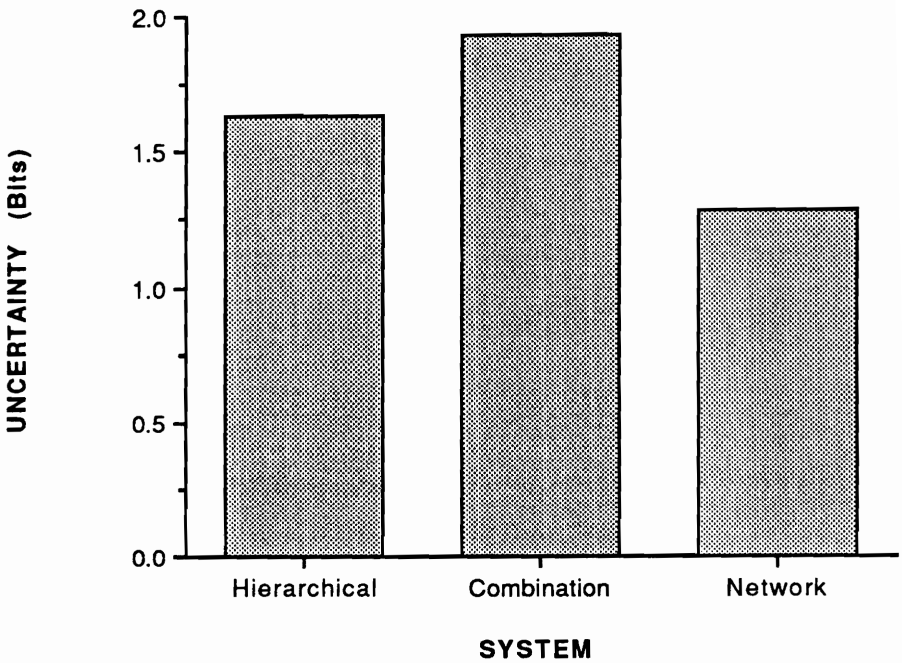


Figure 46. System main effect for uncertainty in Experiment B.

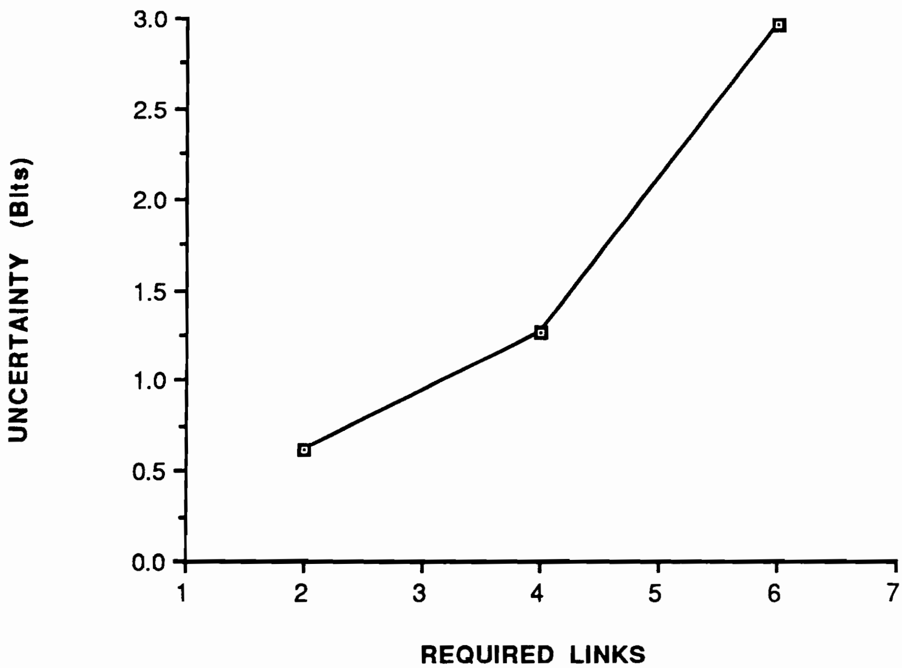


Figure 47. Required links effect for uncertainty in Experiment B.

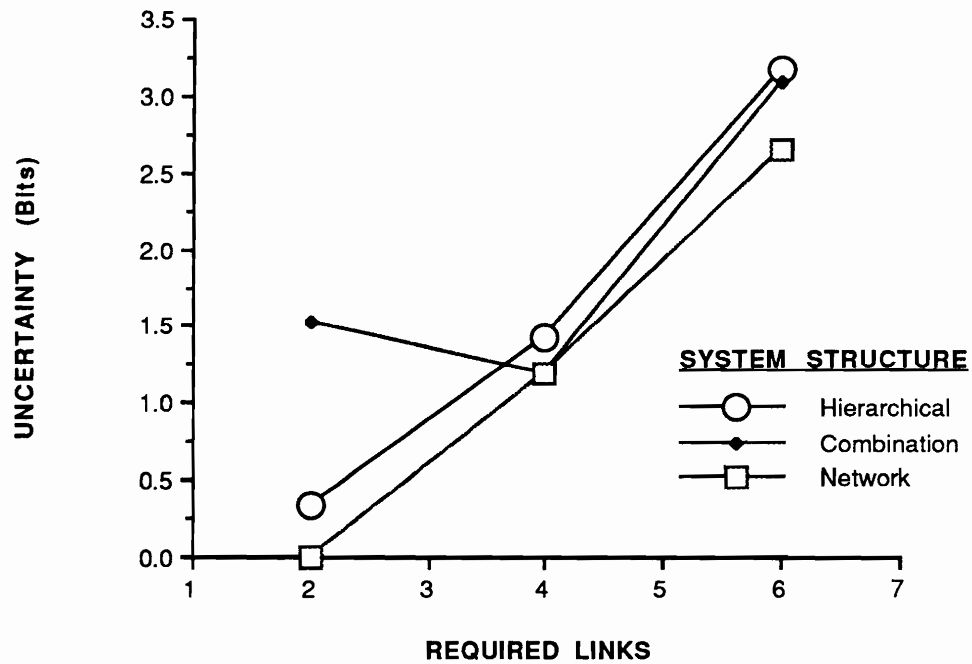


Figure 48. System x Links interaction for uncertainty in Experiment B. With the exception of the 2-link task in the combination condition, the general patterns are similar.

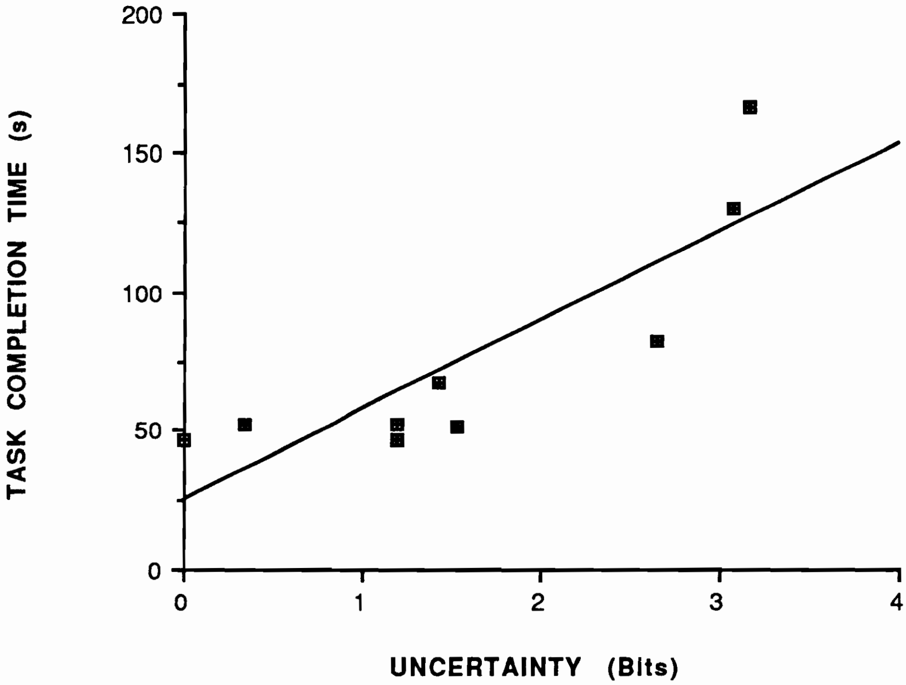


Figure 49. Task completion times plotted against associated uncertainty for each task performed in Experiment B across systems ($r = 0.848$).

TABLE 17. ANOVA Summary Table for Deviations (Experiment B). The included *p*-values have been corrected by the Greenhouse-Geisser (1959) ε where necessary.

SOURCE OF VARIANCE	df	SS	MS	F	p
Between-Subjects					
System (S)	2	8.661	4.331	5.75	0.0060
Subjects/System (S/S)	45	33.916	0.754		
Within-Subjects					
Links (L)	2	20.534	10.267	14.03	0.0001
L x S	4	11.0673	2.767	3.78	0.0112
L x S/S	90	65.848	0.732		
Total	143	140.026			

Greenhouse-Geisser Epsilon value for L and L x S = 0.8343

Both System Structure and Required Links produced significant main effects. Figure 50 presents the System effect for deviation. A Newman-Keuls post-hoc test (Table 18) failed to detect a difference between the network and hierarchical systems or between hierarchical and combination systems. Figure 51 depicts the required links main effect; 6-link tasks have significantly higher deviation than 2- and 4-link tasks, which show no difference. The System x Links interaction (Figure 52) yielded no differences among the systems at 2- or 4-link tasks. At 6-link tasks the combination and hierarchical systems produce significantly higher deviation (Table 19). The combination system at 2-links is not significantly different from the combination and hierarchical configurations at 6-links.

Time Per Screen (Node). A two-way ANOVA analyzed the time per screen data. As seen in Table 20, Required Links, System Structure, and System X Links produced significant results; Figures 53, 54, 55 present these effects. A Newman-Keuls post-hoc test (overall $\alpha = 0.05$) of the number of required links uncovered significantly longer times per screen for 2-link tasks than for 4- or 6-link tasks, which exhibited no difference. Analysis of the System main effect found the hierarchical system to have the shortest times per screen with no difference between the combination and network systems (Table 21). This result is consistent with its counterpart from Experiment A and is further explained by the System x Links interaction. At each level of required links the hierarchical system yields significantly shorter times than the network or combination systems (Table 22). No differences were detected between the network and combination systems for 4- and 6-link tasks.

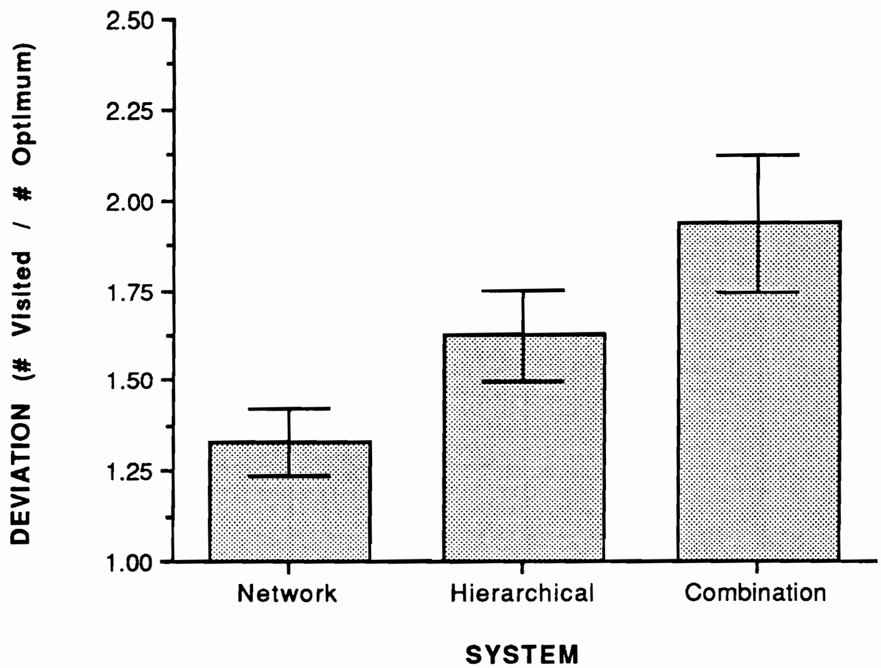


Figure 50. System effect for deviation in Experiment B.

TABLE 18. Newman-Keuls Post Hoc Test for Deviaton System Main Effect ($p < 0.05$) in Experiment B. Deviations with identical letters are not significantly different.

SYSTEM	DEVIATION		
Network	1.330	A	
Hierarchical	1.626	A	B
Combination	1.931		B

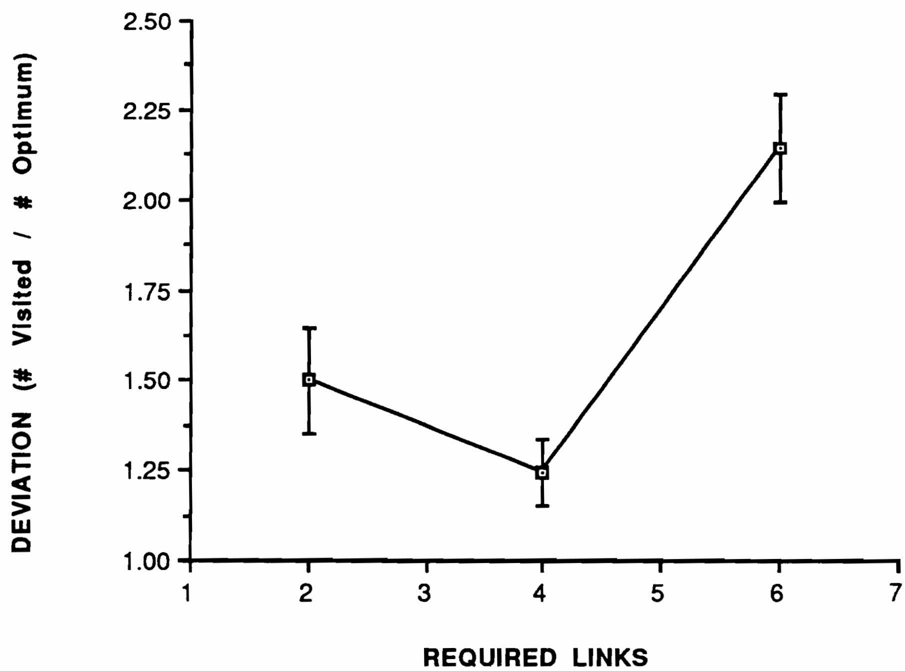


Figure 51. Required links effect for deviation in Experiment B.

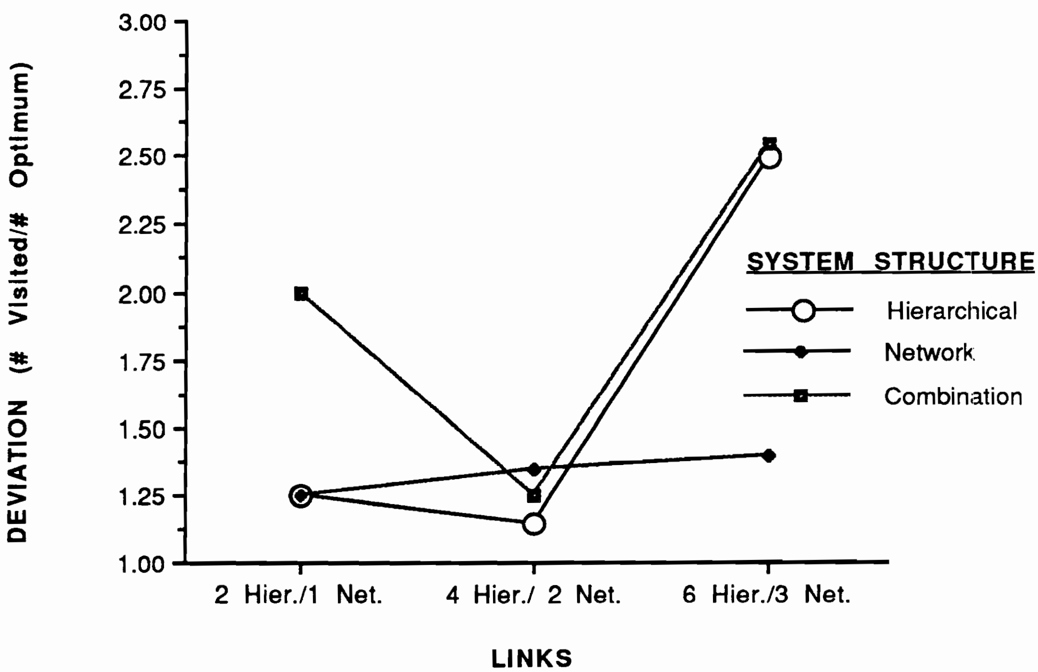


Figure 52. System x Link interaction for deviation in Experiment B. A Newman-Keuls post-hoc test found no difference among the systems for 2-link tasks; however, the hierarchical and combination systems were significantly worse than the network system for 6-link tasks (overall a = 0.05).

TABLE 19. Newman-Keuls Post Hoc Test for Deviation System x Required Links Interaction ($p < 0.05$) in Experiment B. Deviations with identical letters are not significantly different.

SYSTEM	REQUIRED LINKS	DEVIATIONS		
Hierarchical	4-link	1.141	A	
Hierarchical	2-link	1.250	A	
Network	2-link (i.e., 1-link)	1.250	A	
Combination	4-link	1.250	A	
Network	4-link (i.e., 2-link)	1.344	A	
Network	6-link (i.e., 3-link)	1.396	A	
Combination	2-link	2.000	A	B
Hierarchical	6-link	2.490		B
Combination	6-link	2.542		B

TABLE 20. ANOVA Summary Table for Times Per Screen (Experiment B). The included *p*-values have been corrected by the Greenhouse-Geisser (1959) ε where necessary.

SOURCE OF VARIANCE	df	SS	MS	F	p
Between-Subjects					
System (S)	2	2919.417	1459.709	8.47	0.0008
Subjects/System (S/S)	45	7758.207	172.405		
Within-Subjects					
Links (L)	2	4610.329	2305.165	39.46	0.0001
L x S	4	770.796	192.699	3.30	0.0233
L x S/S	90	5257.995	58.422		
Total	143	21316.74			

Greenhouse-Geisser Epsilon value for L and L x S = 0.8077

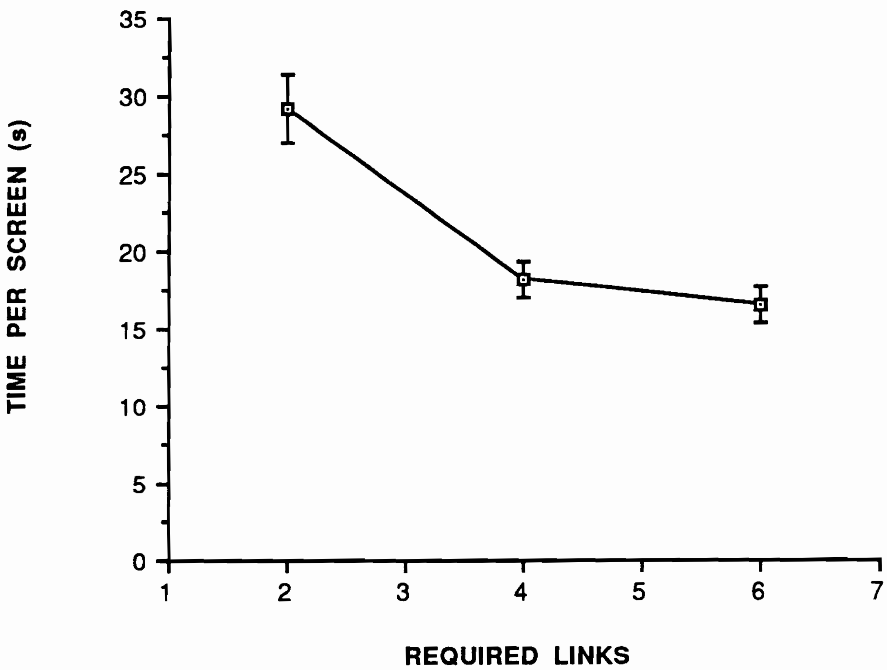


Figure 53. Required links main effect for time per screen in Experiment B.

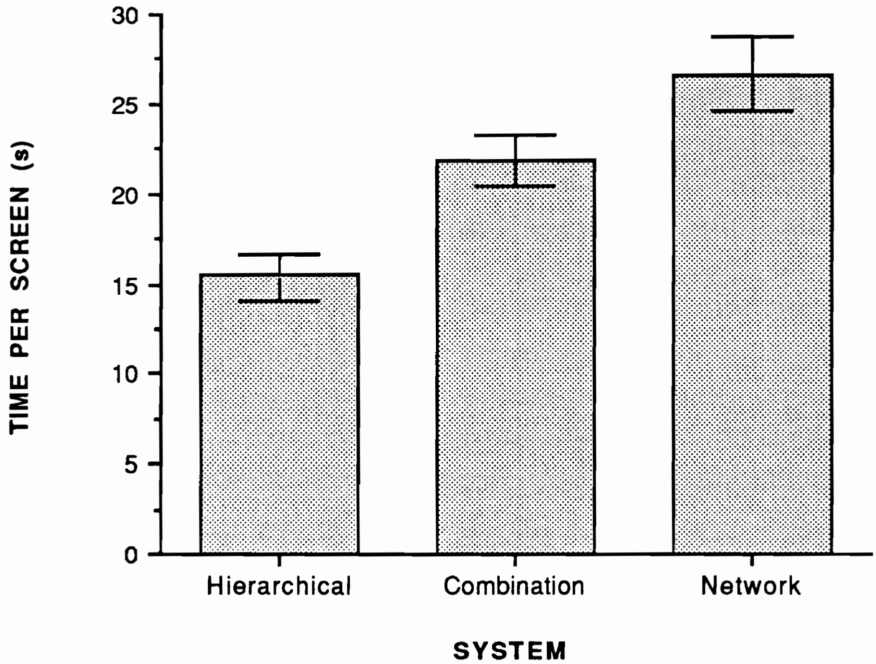


Figure 54. System effect for time per screen in Experiment B.

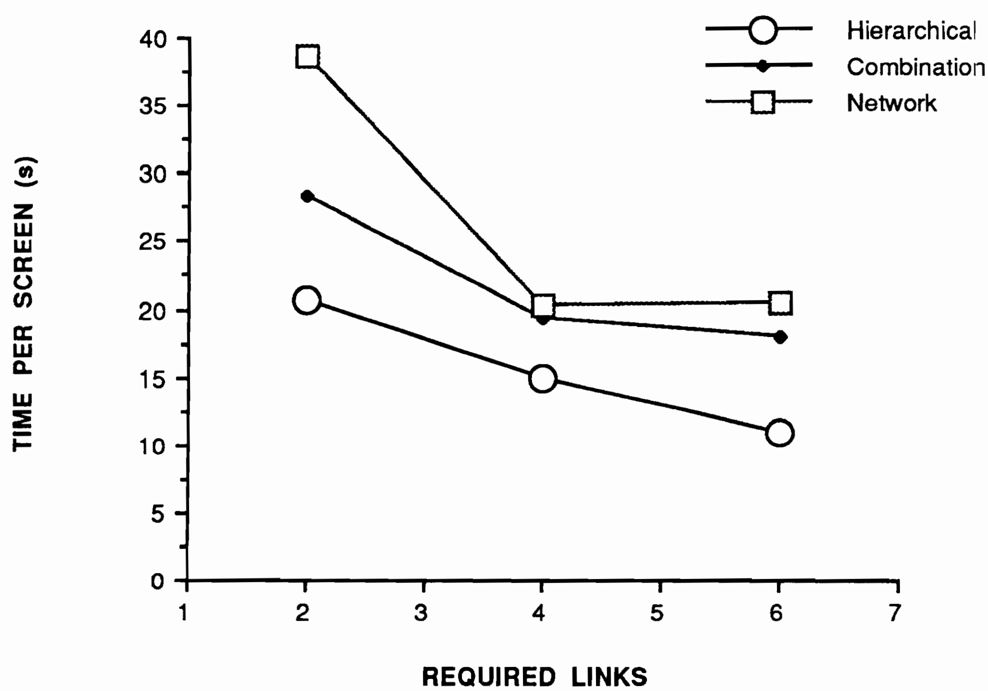


Figure 55. System x Link interaction for time per screen in Experiment B.

TABLE 21. Newman-Keuls Post Hoc Test for Time per Node System Main Effect ($p < 0.05$) in Experiment B. Times with identical letters are not significantly different.

SYSTEM	TIME/NODE (s)	
Hierarchical	15.507	A
Combination	21.867	B
Network	26.491	B

TABLE 22. Newman-Keuls Post Hoc Test for Time per Node System x Required Links Interaction ($p < 0.05$) in Experiment B. Times with identical letters are not significantly different.

SYSTEM	REQUIRED LINKS	TIME/NODE (s)		
Hierarchical	6-link	10.864	A	
Hierarchical	4-link	15.023	A	B
Combination	6-link	17.950	C	B
Combination	4-link	19.349	C	
Network	4-link (i.e., 2-link)	20.210	C	
Network	6-link (i.e., 3-link)	20.519	C	
Hierarchical	2-link	20.640	C	
Combination	2-link	28.304	D	
Network	2-link (i.e., 1-link)	38.740	E	

Subjective Measures (Experiments A and B)

Previous research (Bowers, 1990; Mohageg, 1989) has shown that quick, error-free task performance yields positive subjective ratings. This trend was anticipated for the current study as well; however, only partially similar results were obtained. The modified QUIS was analyzed using a one-factor MANOVA (System being the only factor); resulting parameters for each questionnaire item are included in Appendix VII. Questions 1, 3, 4, 5, 6, 11, 22, 27, 28, 29, and 30 produced significant effects. Table 23 lists the items and the ratings associated with each system.

Questions 1, 3, 4, 5, and 6 solicited overall reactions to the software. Newman-Keuls post-hoc tests failed to uncover differences among the hierarchical, network, and combination systems; however, the linear system was rated significantly lower on each of these five items. Thus, participants found the linear system especially poor. Question 11 addressed the issue of consistency of terms in the system. The network system was rated significantly lower than the remaining three systems, which failed to produce significant differences. This result seems reasonable since there is variability between nodes in the network system with respect to linked terms. Certainly, the network system offers more node-to-node variation than do the hierarchical and linear systems. Many consider this variability to be "flexibility" and deem it necessary for useful hypertext / hypermedia systems.

Question 22 found the linear system rated as significantly slower (system speed) than the other systems, which exhibited no significant differences.

TABLE 23. Obtained Ratings for the QUIS Items. Questions 1, 3, 4, 5, 6, 11, 22, 27, 28, 29, and 30 produced significant differences.

QUESTIONNAIRE ITEM	SYSTEM STRUCTURE Hier.	Net.	H+N	Lin.
1. Overall the system was terrible--wonderful	7.1	7.6	7.2	5.6
2. Overall the system was difficult--easy	7.4	7.0	7.5	6.5
3. Overall the system was frustrating--satisfying	6.8	7.3	7.4	5.6
4. System has inadequate power--adequate power	7.6	7.4	7.7	5.1
5. Overall the system was dull--stimulating	6.8	6.9	6.6	5.1
6. Overall the system was rigid--flexible	6.6	6.8	7.0	5.1
7. Characters hard to read--easy to read	7.3	6.9	7.0	7.4
8. Highlighting simplifies task none--very much	Subjects provided no rating			
9. Organization of information confusing--clear	7.3	7.0	7.6	7.0
10. Sequence of screens is confusing--clear	7.5	6.8	7.0	7.0
11. Term usage was inconsistent--consistent	8.1	7.4	8.4	8.4
12. Terminology is never--always related to task	5.5	5.2	4.1	5.5
13. Position of messages is inconsistent--consiste	4.6	3.3	3.6	4.4
14. Computer keeps you informed	7.3	6.6	6.5	6.0
15. Error messages were unhelpful--helpful	4.0	3.2	3.9	4.0
16. Learning to operate was difficult--easy	8.1	8.3	8.4	7.5
17. Exploring new features was difficult--easy	7.6	7.3	8.3	7.3
18. Recall and use of commands was difficult--easy	7.9	7.3	8.1	6.9

TABLE 23 (Cont.). QUIS Items.

QUESTIONNAIRE ITEM	SYSTEM STRUCTURE Hier.	Net.	H+N	Lin.
19. Tasks never--always done straight-fowardly	7.4	6.9	7.5	6.6
20. Help messages were unhelpful--helpful	2.3	1.8	2.6	1.5
21. Reference materials were unhelpful--helpful	3.7	2.3	3.0	1.8
22. System speed was too slow--fast enough	8.1	7.6	7.6	4.6
23. System was unreliable--reliable	7.3	7.7	7.9	7.4
24. System tended to be noisy--quiet	8.3	8.3	8.4	7.9
25. Recovering from errors was difficult--easy	7.2	7.8	7.8	6.6
26. Expert/novice needs never--always considered	7.3	6.6	7.2	6.8
27. Perfroming tasks was difficult--easy	7.7	7.7	8.1	6.7
28. Getting to desired screen was confusing--clear	7.8	7.6	8.1	6.6
29. You lost your place--never lost place	7.6	7.4	7.6	6.4
30. Provide overall rating from 1 (low) to 9 (high)	8.1	6.2	8.2	4.8

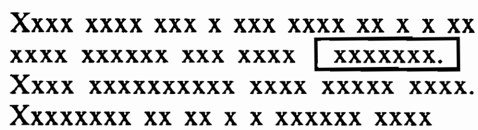
Actual system response times were equivalent for all systems; however, the linear system was especially cumbersome to use. The result found in question 22 is believed to reflect subjects' perceptions of task completion time as influenced by the cumber associated with this system.

Finally, questions 27, 28, and 29 solicited ratings of specific task performance issues, while question 30 asked for an overall preference rating (of 1 to 9). For ease of performing tasks, the combination system was rated as significantly easier than the linear system with no differences existing among other paired systems. The identical result was obtained for accessing the desired system node (question 28). For maintaining a sense of orientation, participants found the linear and network systems to be no different, but rated them significantly lower than the hierarchical and combination systems (which show no reliable difference). Finally, the overall system ratings found the linear system rated reliably lower than the network system, both of which were rated lower than the non-differing hierarchical and combination systems.

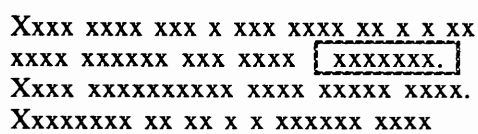
DISCUSSION AND CONCLUSIONS

Task Completion Time (TCT)

The hierarchical and combination systems produced significantly shorter TCTs than the linear and network systems in Experiment A. These systems were identical in node format and content, the chief differences being the linking structures and the presence of marked (cued) words embedded in the text for the network and combination systems (Figures 16 and 17). Thus, the contribution to observed TCT differences of each these two components should be investigated; it is reasonable to postulate a detrimental influence on reading speed due to embedded marked words. Watkins (1990) investigated the influence of marked words on reading speeds for Tinker passages. Of the six cueing configurations used by Watkins, one in particular was quite similar to that used in the current study (Figure 56); data pertaining to this cue were evaluated to determine the influence on reading speed. The Watkins study included 7, 5, 3, and 0 marked words per Tinker passage (roughly 30-35 words). Each subject received all marked word conditions. Resulting data are presented in Figure 57. Six multiple paired comparisons using a t-test (total $\alpha = 0.05$, individual $p < 0.0085$) found significantly longer reading times for 7 marked words than for 0 ($t = 4.055$, $p = 0.0001$), but all remaining comparisons failed to yield significant differences. Of particular import to the current study is the difference between 0 marked words and 1 marked word per passage: nodes in the GEO system (network and combination configurations) averaged



(a)



(b)

Figure 56. Illustration of the mark used to indicate linked words in the GEO system (a), and the box-type mark used in Tinker passages by Watkins (1990) (b).

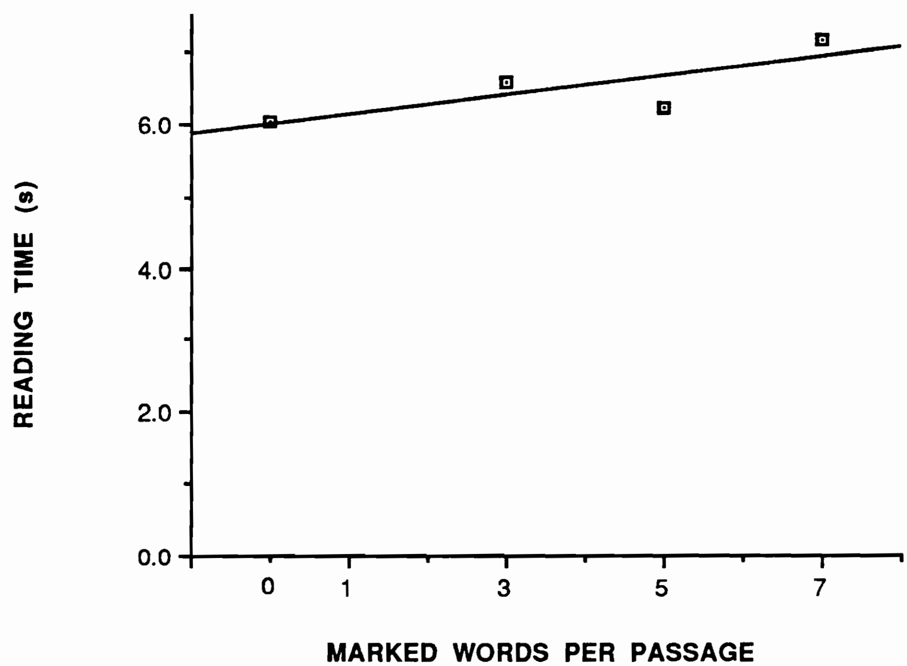


Figure 57. Regression line fit to mean Tinker passage reading times obtained by Watkins (1990). The regression equation was *Time*= 5.9961 + 0.129*Words*, $r^2 = 0.623$.

between 140-180 words with 5 marked terms per node, yielding roughly 1 marked word per 30 words (1 per Tinker passage). The absence of differences among 5, 3, and 0 marked words are strong indication that no differences would exist between 0 and 1; in addition, simple visual inspection of Figure 57 suggests the difference between 0 and 1 is negligible. Therefore, the observed TCT differences among the systems can be confidently attributed to differences in linking structures alone. The System x Links interaction reveals that the TCT advantages of the hierarchical and combination systems are due mainly to their superior performance for 4- and 6-link tasks. The problematic performance of the network system is especially acute for 6-link tasks. As indicated by significant differences in the slopes of the network and hierarchical systems, link traversals of larger than 6 will certainly exacerbate the problematic components of the network structure.

Of particular interest in Experiment A was the performance evaluation of the hierarchical, network, and combination systems for different levels of task type. The System x Task Type interaction reiterated the TCT advantages of the hierarchical and combination systems. For relational tasks, the network system, which was judged to be well suited to these tasks, offered no TCT advantage over the hierarchical configuration and performed significantly worse than the combination system. For hierarchical tasks, the network system proved detrimental to task performance since it yielded TCTs significantly longer than the hierarchical and combination systems. No difference exists between hierarchical and combination structures for either task type.

Hence, the hierarchical linking structure is superior to the network structure where identical numbers of links are required to achieve answers. The hierarchical linking structure is believed to be superior due to its inherent lack of flexibility and sufficient navigational predictability. Searching the hierarchical database is structure-driven; since available links were identical at similar levels in the hierarchy users could formulate structure-based strategies for their search. Most subjects commented positively on the convenience of structure predictability. Conversely, network links were established purely on informational bases (the keyword search) and forced the user into an information-driven search pattern. Available links at each node are highly variable since the nature of the information and corresponding links are inconsistent from node to node. This inconsistency was apparent to the subjects as evidenced by the response to question 11 of the post-experiment questionnaire, which found the network system rated significantly lower than others for consistency of terminology. Subject task performance indicates IR tasks are better performed using a structure-driven approach rather than an information-driven one. Further evidence of advantageous hierarchical linking is uncovered in reviewing the link selection patterns of subjects using the combination system. These subjects traversed a total of 529 hierarchical and relational links, excluding standard links such as "previous screen" and "home." Significantly more hierarchical links were traversed (296) than relational links (233), $z = 2.78$, $p = 0.0054$.

In Experiment B, the network configuration produced reliably shorter TCTs than the linear and hierarchical systems, but exhibited no TCT advantages over the combination system. The System x Links interaction suggests the superior performance of the network system is attributable to 6-link tasks (requiring only three links on the network system). At 2- and 4-link tasks the hierarchical configuration is no different from the network system even though, under optimal link traversals, the network configuration required half as many links (i.e., 1 and 2 links, respectively). The network configuration eventually shows benefits where required number of links were significantly reduced (6 hierarchical vs. 3 network). The combination system showed advantages over the hierarchical configuration for 6-link tasks (2- and 4-link tasks failed to produce a difference), but exhibits reliably longer TCTs than the network configuration. No significant differences not exist at the $\alpha = 0.05$ level between traversed links of the hierarchical (109) and relational (89) varieties ($z = 1.49$, $p = 0.1362$). The combination system shows no advantages over the hierarchical system in Experiment A; however, a slight advantage is detected in Experiment B for 6-link tasks.

All Experiment B tasks were relational tasks, as determined by judges, and required half as many links with the network system. Thus, the task situations heavily favored the network system. Yet, only 6-link tasks (3-link in network system) yield significant differences in TCT between the hierarchical and network configurations. These results indicate there are no TCT advantages in network linking unless the linked terms are well suited to task requirements **and** these links reduce the overall number of link traversals considerably (i.e., from 6 to 3). Moreover, no differences were found between tasks requiring 4 links on the hierarchical system and those requiring 3 links (6 in hierarchical) in the network system, illustrating that the network system can provide TCT advantages only with considerable reduction in required links. These results reconfirm those of Experiment A, which found the hierarchical linking structure to be a superior configuration for IR tasks as measured by TCT.

Uncertainty

Results from the uncertainty measures are consistent with TCT in Experiment A; in fact, uncertainty is highly correlated with TCT ($r = 0.858$), indicating the sensitivity of this newly applied measure to the independent variables. The network condition exhibits higher uncertainty than both the hierarchical and combination conditions. This difference is in great measure due to the exceptionally poor performance of the network system for 6-link tasks. The information theoretical formulation for calculating uncertainty was found to be an effective and viable mechanism for measuring navigational difficulty. The assumption is that if the route for achieving the answer is obvious to users uncertainty will be low. Systems with unclear, illogical, or otherwise not easily uncovered routes will have higher associated uncertainty values. Considerably more varied paths were used during IR task performance in the network system than in the hierarchical or combination systems in Experiment A. The consistent structure of hierarchical linking allowed for less erratic navigation through the system, whereas, the unpredictability of the network system led to nonuniform navigation patterns. It is interesting to note that the uncertainty value is not simply a function of the number of available links (i.e., more links lead to higher uncertainty). The combination system offered twice as many links at each node than did the network or hierarchical links, but still produced fairly low uncertainty in Experiment A.

While TCT and uncertainty are well correlated in Experiment B as well (r = 0.848), no reliable difference exists between the hierarchical and network systems. This absence of a difference is consistent across number of required links. The combination system exhibits unusually high uncertainty for the 2-link task; this result is believed to be an uncharacteristic occurrence. Again, the advantages of the hierarchical configuration are reaffirmed: the network system fails to show lower uncertainty in spite of requiring half as many links, and the combination condition provides no advantages.

Deviations from the Optimum Path

The exceptionally high deviation uncovered for the hierarchical system was due to relational tasks in Experiment A (significant System x Task Type interaction). When considered in combination with TCT and uncertainty results, this finding is quite surprising. Subjects travel to significantly more (and perhaps unnecessary) nodes in the hierarchical configuration with very high speed, as evidenced by the short TCTs, and deviate to the identical nodes, as indicated by low uncertainty. Identical deviations seem particularly suspect since the possibility for multiple variations is considerably higher than similar variations. A review of search tendencies for performing relational tasks uncovered an interesting navigational pattern. For the 6-link relational task (task 7), subjects were required to access information from two separate nodes to complete the task satisfactorily. Both target nodes were leaf nodes within two chapters; thus, users were required to travel down the hierarchy in one chapter, return to the home node, and navigate down the hierarchy in the second chapter to reach the final target node. In defining the optimum path for this task, the experimenter assumed use of the "Home" button to directly access the top level node in the hierarchy. This single link eliminates the need for retracing 3 steps upwardly in the hierarchy. However, 44% of subjects failed to make use of the "Home" button, opting instead to retrace their steps in the hierarchy. This navigational behavior is due to either the subjects' poor understanding of the "Home" button or their preference for retracing steps to maintain a sense of orientation. The author believes the latter explanation to be more viable since other subjects using the hierarchical structure and exposed to identical training materials effectively utilized the "Home" button. Individual differences (e.g., spatial ability, orientation ability) may play a role in determining navigational behavior; however, it is important to note that subjects choosing to retrace their steps (perhaps due to individual differences) did not significantly impact TCT or uncertainty, indicating the task performance superiority of the hierarchical configuration.

A second navigational pattern contributing to the discussed result was that many subjects traveled down the incorrect section tree once entering the chapter (country) of interest. Six of 16 subjects (38%) chose to travel to "Economy of Morocco" rather than "Geography of Morocco" for IR task 7. Recovering from this error was performed identically by all six subjects. Again, this pattern increased the deviation without considerably impacting the uncertainty.

In Experiment B, no overall difference was found between the network and hierarchical configurations or between the hierarchical and combination systems. The network system exhibits less deviation than the combination, mainly due to its advantages for 6-link tasks. Additionally, it seems users are unable to usefully employ network links in the combination system for tasks requiring numerous traversals. For the 6-link tasks only significantly fewer network links (45) were traversed than hierarchical links (82), $z = 3.37$, $p = 0.001$. This may indicate that users experiment with network links only when they perceive minimal risk, as in situations requiring few links. However, users may be reluctant to use unpredictable network links when significant navigation is required (or if they believe such to be the case).

Time Per Node (Screen)

The significantly shorter times per node for the hierarchical and combination structures are attributed to the navigational ease provided by hierarchical linking. The mean difference between the hierarchical and network systems was roughly 9 s per node.

Times per node provided further evidence of the task-relevant advantages offered by the structure-based configuration of hierarchical links. Subjects were able to quickly develop a browsing strategy, use the memorable structure to reach the potential target node, and scan the node for appropriate information. Hence, evaluation of each link of a given node was not necessary since subjects were fairly certain of the options they would encounter at each node. Link predictability was true of one-half the available links in the combination structure: the hierarchical links. The time per node was not significantly higher for this system than the hierarchy alone, perhaps due to available and highly used hierarchical links (indicated by the frequent use of hierarchical links in the combination system).

The situation is quite different for the network structure. Information-based linking led to considerable inconsistency among nodes. Upon arriving at a given node, subjects were forced to evaluate each link and determine its relevance to the task at hand. A particularly important consequence of this link variation is the subjects' inability to formulate a reliable strategy before browsing for the information; a novel and perhaps fluid strategy was adopted for each task based upon accessed nodes, the available links at each node, and their level of applicability to the task. Such a strategy would undeniably lead to considerably longer time spent at each encountered node. Given that tasks required the access of at least the same number of nodes (as determined by number of required links), it is reasonable that final TCTs were significantly shorter for the hierarchical and combination conditions.

The hierarchical system exhibited significantly lower times per screen across all levels of required links in Experiment B. While minimal time spent per node is not an inherently valuable advantage, it does indicate a significant amount of user confidence in moving quickly through nodes; whereas, users of the network condition seemed more conservative in travelling from node to node.

Subjective Measures

In general, the subjective measures illustrate the perceived deficiencies of the linear system. The poor ratings are not surprising since the cumbersomeness associated with the linear system can easily lead to frustration and disgruntlement, resulting in low ratings.

Methodological considerations may have removed any possibility of obtaining differential ratings among the systems. The concern of transfer of training necessitated system linking structure to be a between-subjects variable. The systems and information domain were novel to subjects and most enjoyed the experience of using the systems (especially the hierarchical, combination, and network systems). For the most part subjects claimed to have pleasant interactions with the systems, resulting in consistently high ratings for a majority of questionnaire items. A relative comparison may have been more powerful and, perhaps, more meaningful than the obtained absolute ratings; however, such a comparison would require interaction with all systems.

Conclusions

Linear linking. It is apparent that the linear linking structure is a less efficient alternative when compared to the other systems. This difference is especially acute in comparison to the combination and hierarchical configurations. The linear structure included in this study lacks highly sophisticated navigation features. For instance, each node of the system has an entry in the table of contents, but users cannot click on the entry and access the node. While many systems provide such options, the intent of the current study was to measure performance given usage of specific linking structures. The provision of immediate access to nodes through a table of contents would defeat this purpose. Additionally, specific node titles may not always correspond to IR task requirements; hence, users would be required to enter the system and browse about the document in search of the target information. Therefore, while the linear configuration was not augmented with sophisticated searching mechanisms (neither were the experimental systems), the basic linear organization of nodes, as seen in many systems, was maintained.

Hierarchical linking. The multiple advantages of the hierarchical configuration suggest hierarchical linking structures should be implemented for most hypertext (or text-intensive) databases used for IR. The taxonomical issue of whether to label hierarchically organized systems as "hypertext" is certainly minimized (and perhaps irrelevant) since task performance issues of such systems are more salient concerns. Thus, regardless of accepted nomenclature, the hierarchical organization is highly useful in organizing text-intensive information for IR. This recommendation does not exclude the use of other organizational structures, but does indicate the necessity of the hierarchy as an underlying baseline organizing structure. While results of the current study are most directly applicable to IR situations, it is believed that the advantages of hierarchical linking transcend application type and should be of benefit in a majority of hypertext applications (e.g., education, entertainment, etc.).

Network linking. Implementation of network links was based solely in keyword searches. While this method of creating relational links is practiced often, it is not the only strategy for creating such links. The results from this study are intended to generalize to such relational links.

Results of the current study imply the use of network (relational) links in isolation from organizational linking should be strongly discouraged. Their inclusion should be in combination with other types of linking; the hierarchical organization proved a suitable structure for such organizational linking.

Additionally, care should be taken in the establishment of network links. A critical issue will be for designers to identify task-relevant keywords since network links not suited to task requirements prove detrimental, not harmless. Moreover, relevant network links capable of saving users many steps (or traversed links) can be beneficial.

The extra costs of network linking, both during system development and in requiring system resources, are not justified by the performance gains for IR tasks. In fact, under certain conditions network links show negative impact on such task performance. These statements, however, must be interpreted in light of the subject population of the current study; subjects were novices to both the information domain and the GEO system. The utility of network linking may be higher for domain experts or experienced system users. Nonetheless, network linking in isolation is unable to provide practical gains for novice system users.

Combination linking. In terms of task completion time only, the combination system performed as well as the best system in each experiment: no difference exists between combination and hierarchical configurations for Experiment A nor between combination and network configurations for Experiment B. Thus, if task completion time were the most salient consideration the combination system would be the configuration of choice. While the data fail to overwhelmingly support the use of combination linking, it is believed that the combination configuration is no worse than the best systems and may be advantageous over time. That is, while no differences may exist for a few tasks, over a period of performing many tasks the access to both types of linking may be beneficial, especially for task completion time. Ultimately, the decision to include both hierarchical and network links is likely to be determined by the information domain and the designers' ability to identify useful keywords. This author recommends hierarchical organization of material as the emphasized and underlying structure in addition to which network links are included on a secondary (less important) level.

Future Research

While the current study includes previously uninvestigated variables, there is a myriad of variables remaining to be studied. Empirical investigations of factors influencing hypertext/hypermedia use will undeniably lead to more efficient and prudent applications of these systems.

Information domain experts and system experts. User knowledge is perhaps the most salient mediating variable not included in this study. Both information domain and system knowledge may significantly influence the utility and viability of hypertext systems, in particular relational linking. Figure 58 illustrates the "experience matrix" and the matrix location of subjects in the current study. Obviously, subjects were novices in the information domain and with regard to the system. Conceivably, users could learn the underlying hierarchical structure after considerable exposure. Subsequent to this exposure users may feel more at ease and confident in using relational links to navigate within the document. Additionally, domain experts may immediately find utility in relational linking. Linked words often lead to desired information and domain experts may recognize this fact, whereas novice users may fail to recognize the utility or importance of a particular link. A combination of high

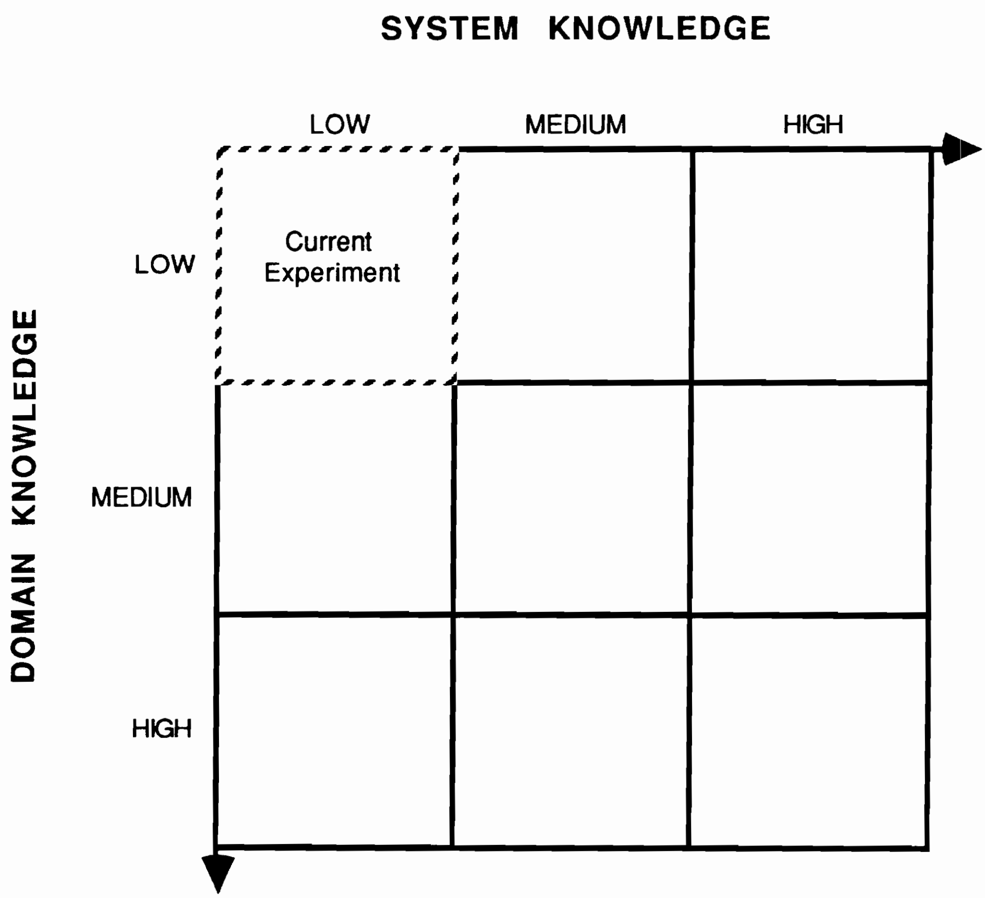


Figure 58. The "Experience Matrix." Arrows are included to indicate these axes are continuous; the delineated categories are arbitrarily defined. Subjects participating in the current experiment fall in the low category on both dimensions. Various combinations of system and domain knowledge may have a mediating influence on hypertext usability.

system knowledge and high domain knowledge may yield interesting results, which could augment those of the current study.

Alternative relational link generation techniques. The method of link generation used in the current study is not the only technique for such linking. As discussed in the Literature Review, there is a variety of link generation techniques. Future research should focus on these alternative methods and evaluate their relative merits in addition to comparisons with standard organizational linking. The keyword-based link generation technique used in GEO is certainly representative of common practice; however, common practice does not necessarily define the optimum.

Research on larger systems. Hypertext system size is an additional important issue. The tradeoffs uncovered in the current study may be different for larger systems. For instance, in a very deep hierarchy, hierarchical links may be quite cumbersome; however, a large system also offers more opportunity for disorientation with relational links. Maps and other navigational guides will be important considerations in the investigation of large systems.

User-defined linking. As hypertext systems mature the option of user-defined linking will gain relevance. Designer-established links will fail to satisfy all IR situations and the availability of user-defined linking may become essential. Therefore, issues concerning user-established links will have considerable impact on hypertext usability. Two important considerations will be (1) Will users be willing to expend the time and effort required to establish links?, and (2) will most user-established links be hierarchical? The costs associated with link creation may not appeal to many users, in which case user interface designers must provide capabilities to allow for fairly effortless link creation. In addition, most users may inherently prefer hierarchies, as the results indicate, and use such linking by default. Again, mechanisms for easy creation and management of such links should be considered.

Integration of browsing and other search mechanisms. The focus of the current study is browsing through a database in search of specific information. A multitude of additional mechanisms has been used to perform similar tasks or, as discussed earlier, narrow the focus browsing. Brown (1988) has discussed the issue of satisfactorily combining search mechanisms in a hypertext system. While hypertext systems are useful and operable with browsing, string search or query language techniques are also helpful. Therefore, combining these search mechanisms in a complementary and useful fashion may be critical to system usability.

REFERENCES

Akscyn, R. M., McCracken, D. L., and Yoder, E. A. (1988). KMS: A distributed hypermedia system for managing knowledge in organizations. *Communications of the ACM, 31* (7), 820-835.

Allen, R. B. (1983). Cognitive factors in the use of menus and trees: An experiment. *IEEE Journal on Selected Areas in Communication, SAC-1* (2), 333-336.

Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods, and results.* New York: Henry Holt and Company.

Belkin, N. J. and Croft, W. B. (1987). Retrieval techniques. In M. E. Williams (ed.) *Annual review of information science and technology (ARIST), Vol. 22.* Amsterdam: Elsevier Science Publishers B. V.

Bennett, J. L. (1984). Managing to meet usability requirements. In J. Bennett, D. Case, and J. Sandelin (Eds.), *Visual display terminals* (pp. 161-184). Englewood Cliffs, NJ: Prentice-Hall.

Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology, 1*, 18-46

Bowers (1990). Concurrent vs. retrospective verbal protocol for comparing window usability. Unpublished Doctoral dissertation. Virginia Polytechnic Institute and State University: Blacksburg, VA.

Brown, P. J. (1988). Linking and searching within hypertext. *Electronic Publishing, 1* (1), 45-53.

Bush, V. (1967). *Science is not enough.* New York: William Morrow and Company, Inc.

Campagnoni, F. R., and Ehrlich, K. (1989). Information retrieval using a hypertext-based help system. Paper presented at the 12th International Conference on Research and Development in Information Retrieval, Cambridge, Mass.

Cecala, T. Watkins, R. and Mohageg, M. F. (1989). *Human computer research annual technical report.* (Technical Report No. CSC-TR89-014). Texas Instruments, Inc., Computer Science Center, User Systems Engineering Laboratory, Dallas, TX.

Conklin, J. (1987). *A survey of hypertext* (Report Number STP-356-86). Austin, TX: Microelectronics and Computer Technology Corp.

Dillon, A., Richardson, J., and McKnight C. (1989). Human factors of journal usage and design of electronic texts. *Interacting with Computers, 1* (2), 183-189.

Dumais, S. T. (1989). Textual information retrieval. In M. Helander (ed.) *Handbook of human-computer interaction.* Amsterdam: Elsevier Science Publishers B. V.

Dumais, S. T., Furnas, G. W., Landauer, T. K., Deerwester, S., and Harshman. R. (1988). Using latent semantic analysis to improve access to textual information. In *CHI '88 Human Factors in Computing Systems*, (pp. 281-285). New York: Association for Computing Machinery.

Egan, D. E., and Gomez, L. M. (1985). Assaying, isolating, and accommodating individual differences in learning a complex skill. In R. Dillon (Ed.), Individual differences in cognition. Vol. 2. New York: Academy Press.

Egan, D. E., Remde, J. R., Landauer, T. K., Lochbaum, C. C., and Gomez, L. M. (1989). Behavioral evaluation and analysis of a hypertext browser. In *CHI '89 Human Factors in Computing Systems*, (pp. 205-210). New York: Association for Computing Machinery.

Fiderio, J. (1988). A grand vision. *BYTE, 9* (10), 237-244.

Frick F. C., and Miller, G. A. (1951). A statistical description of operant conditioning. *American Journal of Psychology, 64* (1), 20-36.

Frisse, M. (1988). From text to hypertext. *BYTE, 9* (10), 247-253.

Furnas G. W. (1986). Generalized fisheye views. In *CHI '86 Human Factors in Computing Systems*, (pp. 16-23). New York: Association for Computing Machinery.

Goodman, D. (1988). *Danny Goodman's HyperCard developer's guide.* New York: Bantam Books.

Gordon, S., Gustavel, J., Moore, J., and Hankey, J. (1988). The effects of hypertext on reader knowledge representation. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 296-300). Santa Monica CA: Human Factors Society.

Gould, J. D., Alfaro, L., Barnes, V., Finn, R., Grischkowsky, N., and Minuto, A. (1987). Reading is slower from CRT displays than from paper: Attempts to isolate a single-variable explanation. *Human Factors, 29* (3), 269-300.

Gould, J. D., Alfaro, L., Finn, R., and Minuto, A. (1987). Reading form CRT displays can be as fast as reading form paper. *Human Factors, 29* (5), 497-517.

Gould, J. D., and Grischkowsky, N. (1984). Doing the same work with hard copy and with cathode-ray tube (CRT) computer terminals. *Human Factors, 26*, 323-337.

Gould, J. D., and Lewis, C. H. (1985). Designing for usability - key principles and what designers think. *Communications of the ACM, 28*, 300-311.

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24* (2), 95-111.

Halasz, F. G. (1988). Reflections on NoteCards: Seven issues for the next generation of hypermedia systems. *Communications of the ACM, 31* (7), 265-283.

Horton, W., Soderston, C., Garstka, J., and Park, M. (1988). *Help and online documentation.* (Available from William Horton and Associates).

Human Factors Society (1988). *American national standard for human engineering of visual display terminal workstations* (ANSI/HFS Standard No. 100-1988). Santa Monica, CA: Human Factors Society, Inc.

Huynh, H., and Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *Journal of the American Statistical Association, 65* (332), 1582-1589.

Jones, W. P. (1987). How do we distinguish the hyper from the hype in non-linear text? In H. J. Bullinger and B. Shackel (Eds.), *Human-Computer Interaction - INTERACT '87* (pp. 1107-1113). Amsterdam: Elsevier Science Publishers B. V.

Jorna, G. C. (1989). The effects of image quality on reading performance and perceived image quality from CRT and hard-copy displays. Unpublished Master's thesis. Virginia Polytechnic Institute and State University: Blacksburg, VA.

Kiger, J. L. (1984). The depth/breadth trade-off in the design of menu-driven user interfaces. *International Journal of Man-Machine Studies, 20*, 201-213.

Lacy, R. M., Chignell, M. H., and Kinnell, S. K. (1988). Authoring hypermedia for computer based instruction. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 313-317). Santa Monica CA: Human Factors Society.

Marchionini, G., and Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer, 21* (1), 70-80.

Mercer, J. (1976). *Spanish Sahara.* London: George Allen & Unmin Ltd.

Miller, D. P. (1981). The depth/breadth tradeoff in hierarchical computer menus. *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 296-300). Santa Monica, CA: Human Factors Society.

Miller, G. (1956). The magic number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Miller, G. A., and Frick, F. C. (1949). Statistical behavioristics and sequences of responses. *Psychological Review, 56* (6), 311-324.

Mohageg, M. F. (1989). Differences in performance and preference for object-oriented vs. bit-mapped graphics interfaces. In *Proceedings of the 33rd Annual Meeting of the Human Factors Society* (385-389). Santa Monica, CA: Human Factors Society.

Monk, A. F., Walsh, P., and Dix, A. J. (1988). A comparison of hypertext, scrolling, and folding as mechanisms for program browsing. In D. M. Jones and R. Winder (Eds.) *People and Computers IV.* Cambridge: Cambridge University Press.

Murray, J. (Ed.). (1981). *Cultural Atlas of Africa.* Oxford: Phaidon Press Ltd.

Muto, W. H. (1990). Personal communication, February 1990.

Nelson, H. D. (Ed.). (1979a). *Libya: A country study.* Washington D.C. : United States Government / American University

Nelson, H. D. (Ed.). (1979b). *Tunisia: A country study.* Washington D.C. : United States Government / American University

Nelson, H. D. (Ed.). (1986). *Algeria: A country study.* Washington D.C. : United States Government / American University

Nelson, T. H. (1967). Getting it out of our system. In G. Scheckter (Ed.), *Information retrieval: A critical review* (pp. 191-210). Washington D. C.: Thompson Books.

Nelson, T. H. (1980). Replacing the printed word: A complete literary system. In S. H. Lavington (ed.), *Information processing '80* (pp. 1013-1023). Amsterdam: North-Holland Publishing Company.

NeXT (1989). *The NeXT computer 0.9/1.0 release description.* NeXT, Inc.

Nielsen, J. (1989). *Introduction to hypertext and hypermedia.* (Available from J. Nielsen, Technical University of Denmark, Dept. of Computer Science, DK-2800 Lyngby Copenhagen, Denmark.)

Nielsen, J. (1990). The art of navigating through hypertext. *Communications of the ACM, 33* (3), 296-310.

Nyrop, R. F. (Ed.). (1983). *Egypt: A country study.* Washington D.C. : United States Government / American University

Parunak, H. V. D. (1989). Hypermedia topologies and user navigation. In *Hypertext '89 Proceedings* (43-50). New York: Association for Computing Machinery.

Raymond, D. R. and Tompa, F. W. (1987). Hypertext and the new Oxford English Dictionary. In *Hypertext '87 Proceedings* (143-153). New York: Association for Computing Machinery.

Reisel, J. F. and Shneiderman B. (1987). Is bigger better? The effects of display size on program reading. In G. Salvendy (ed.), *Social, ergonomic and stress aspects of work with computers* (pp. 113-122). Amsterdam: Elsevier Science Publishers.

Roberts, T. L. and Moran, T. P. (1983). The evaluation of text editors: Methodology and empirical results. *Communications of the ACM, 26* (4), 265-283.

Schwalm, N. D. and Samet, M. G. (1989). Hypermedia: Are we in for "Feature Shock"? *Human Factors Society Bulletin, 32* (6), 1-3.

Shackel, B. (1984). The concept of usability. In J. Bennett, D. Case, and J. Sandelin (Eds.), *Visual display terminals* (pp. 45-87). Englewood Cliffs, NJ: Prentice-Hall.

Shneiderman, B. (1987). *User interface design and evaluation for an electronic encyclopedia* (Report Number CS-TR-1819). College Park, MD: University of Maryland, Department of Computer Science.

Shneiderman, B. (1988). *Reflections on authoring, editing, and managing hypertext*, (Report Number CS-TR-2160). College Park, MD: University of Maryland, Department of Computer Science.

Shneiderman, B. Brethauer, D., Plaisant, C., and Potter, R. (1988). *The HyperTies electronic Encyclopedia: An evaluation based on three museum installations.* (Available from B. Shneiderman, Dept. of Computer Science, University of Maryland, College Park, MD 20742).

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill Book Company.

Smith, E. E., Shoben, E. J., and Rips, L. J. (1974). structure and process in semantic memory: A feature model for semantic decisions. *Psychological Review, 81*, 214-241.

Smith, J. B., and Weiss, S. F. (1987). A hypertext writing environment and its cognitive basis. In *Hypertext '87 Proceedings* (pp. 195-213). New York: Association for Computing Machinery.

Solso, R. L. (1988). *Cognitive psychology.* London: Allyn and Bacon, Inc.

The Cambridge encyclopedia of Africa. (1981) Eds.) Roland Oliver and Michael Crowder. Cambridge: Cambridge University Press.

Tompa, F. W. (1989). A data model for flexible hypertext database systems. *ACM Transactions on Information Systems, 7* (1), 85-100.

Vail, R. (1989). A pilot study of navigational aids' impact on the usability of hierarchical and hypertext databases. In *Proceedings of the 33rd Annual Meeting of the Human Factors Society* (408-412). Santa Monica, CA: Human Factors Society.

Vicente, K., Hayes, B., and Williges, R. C. (1987). Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors, 29* (3), 349-359.

Yankelovich, Meyrowitz, N. K., and van Dam, A. (1985). Reading and writing the electronic book. *IEEE Computer, 18* (10), 15-30.

Yankelovich, N. Haan, B. J., Meyrowitz, N. K., and Drucker, S. M. (1988). Intermedia: The concept and construction of a seamless information environment. *IEEE Computer*, *21* (1), 81-96.

Appendix I. Information Retrieval Tasks.

The following nine tasks were performed, in random order, by the 64 subjects. For each task the number of required links and task type variables are indicated in italics, as is the specific experiment (Experiment A or B).

1. What is the per capita income of the largest nation in North Africa? *(2-link relational task, Experiment A)*

2. School attendance at the elementary level is about 65% in Tunisia. True/False *(2-link hierarchical task, Experiment A)*

3. What percentage of the Egyptian labor force is employed in the agriculture industry? *(2-link relational task, Experiment B; required 1 link on Network system)*

4. The gross national product of Egypt is larger than the gross national product of Libya (if you do not know what the "gross national product" is please feel free to ask). True/False *(4-link hierarchical task, Experiment A)*

5. The year petroleum companies discovered oil in Libya is the same year Libya obtained independence and the right to create a government. True/False *(4-link relational task, Experiment A)*

6. Libyans practice Islam and their population is 30 million people. True/False *(4-link relational task, Experiment B; required 2 links on Network system)*

7. Italy is a major trade partner of Algeria <u>and</u> the capital city of Morocco has an agriculture industry which produces 100% of the city's fruit needs. True/False *(6-link relational task, Experiment A)*

8. In the early 1980s Morocco invaded Western Sahara. Algeria has been supporting the Western Saharan rebels in the Sahara conflict, while Tunisia supports Morocco. True/False *(6-link hierarchical task, Experiment A)*

9. The capital city of Tunisia is more populated than the nation of Algeria. True/False *(6-link relational task, Experiment B; required 3 links on Network system)*

Appendix II. Participant Instructions and Informed Consent Form.

PARTICIPANT'S INSTRUCTIONS

Research Goals

The general goal of this study is to investigate the relative merits of organizing information in specific types of databases to be used in information retrieval. You will perform simple information retrieval tasks in one database, followed by a rating of performing the tasks. Bear in mind that the system is under evaluation, **not you**. Your ability to use the database is heavily influenced by system characteristics; thus, please be candid and frank in your system ratings.

General Overview

The database you are about to use contains information regarding the countries of North Africa. All the information is included in individual screens, which have been connected using machine-supported links between the screens. Your task will be to simply retrieve from the database information that will satisfy different queries or questions. Please attempt to reach the answer as quickly as possible without compromising accuracy. You will be given a brief introduction to the system and will perform three practice trials. A typical task may be to find information confirming or disputing the following statement, "The largest country in North Africa exports 5 agricultural products." As you will see in the introduction, no query language is available; you must use the provided links to navigate about the network and satisfy the task requirements. If you have any questions about the study please ask.

PARTICIPANT'S INFORMED CONSENT FORM

This study intends to evaluate different types of database systems. As a participant, you are helping in the evaluation of the systems. Please keep in mind that the system is under study, **not you**. You have certain rights as explained below. This form will enumerate these rights and obtain your written consent to participate in the study. Your rights as a participant are:

1. You have the right to withdraw from this study at any time and for any reason by simply informing the experimenter. If you decide to terminate your participation, you will be paid only for the amount of time you have participated.
2. You have the right to inspect your data and withdraw them from the experiment if you feel that you should for any reason. Data are processed and analyzed after a subject has completed the study. At that time, these data will be treated with anonymity since all identifying information will be removed from the data. No one else will have access to your data.
3. You have the right to be informed of the overall results of this study (if you so desire). If you wish to receive an overview of the results, include your address with your signature below.
4. If you have any questions prior to data collection please feel free to ask. Unless the outcome of the study will be influenced, the researcher will provide and answer to your satisfaction. Answers which may influence the study will be delayed until after data collection, at which time a full answer will be given.

5. There is minimal risk to you as a participant in this study. Since you are an experienced computer user there is no risk involved in this study beyond what might be experienced in every-day computer usage.
6. You will be paid $ 5/hr. for your time.

If you have further comments or questions about your rights as a participant, please contact Dr. Ernest Stout, chairman of the Institutional Review Board for the Use of Human Subjects in Research. He may be contacted at:

Chairman, University Human Subjects Committee
301 Burruss Hall
Virginia Polytechnic Inst. and State University
Blacksburg, VA 24061
(703) 231-5283

The researcher for this study will be Michael Mohageg, a graduate student in Industrial Engineering and Operations Research.

Your signature below indicates that you have read this document in its entirety, and understand your rights as a participant as listed above, and that you consent to participate. Thanks for your cooperation.

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I so choose.

____

Participant's Signature

____

Printed Name

If you wish to receive a summary of the results print your name and address.

Appendix III. Pre-experiment Questionnaire.

Name: Subject No:

Age:

Sex:

Personal Computer Experience (put a check by the appropriate category):

______Novice

______Have used basic programs (e.g., word processing)

______Fairly experienced user of PCs, have used many packages for a variety of tasks.

______Fairly experienced PC user and have done some simple programming

______Programmer

The following questions are related to the database you are about to use. Please answer these questions to the best of your ability. If you do not know the answer to a question simply leave the answer slot blank.

1. What is the population of Libya? ______________
2. What language do Egyptians speak? ________________
3. What is Morocco's most lucrative export product? ________________
4. What is the capital city of Libya? ______________
5. Mauritania invaded the Western Sahara in the early 1980s. True / False. ___
6. What year did Tunisia gain independence? ________________
7. Who is the current Moroccan king? _________________
8. Cap Bon is in which North African country? ________________

Appendix IV. Training Materials.

THE HIERARCHICAL DATABASE

This database contains roughly 187 screens which have been connected in an hierarchical configuration. The database discusses various issues pertaining to the countries of North Africa. Five major categories exist within each country: demographics, economy, education, geography, and government. No further explanation shall be given regarding the various categories within the hierarchy; as you perform the tasks you will become familiar with the database.

In performing the tasks, your strategy should be to click on the appropriate link indication (located in the upper left-hand corner of the screen) to access the screen which you think may contain the answer to the query. If the answer is not found you should move about the database, using the available links, to reach the screen(s) that might provide the answer. The answer to some queries will require information from only one screen, while others may require information from two screens.

The buttons in the lower left-hand corner of the screen are as follows:

The **Prev. Level** button will take you one level up in the hierarchy.

The **Path History** button will regressively access each screen you have been to. That is, it will move backward one screen-at-a-time. So if you click on this button once you will be taken to the last screen you were at. If you click again you will be taken to your location from 2 screens past, and so on for up to 42 screens.

The **Glossary** button accesses a glossary of link terms. Each of the available link titles (the titles in the buttons) are listed in alphabetical order and a description of the information contained in the referent screen is included.

The **Home** button will take you directly to the top screen in the database (i.e., the North Africa screen).

The **Answer** button is used for indicating the answer. When you believe you've obtained the answer click on this button. An answer screen will appear; simply position the cursor in the appropriate field and type in the answer.

The experimenter will go through the system with you; take 5 minutes to "play around" in the system and familiarize yourself with it. Try clicking on link indications to see where they lead.

Please attempt the following practice queries:

1. The largest country in North Africa exports 5 agricultural products. True or False

2. Wheat and barley constitute as much as 97% of Tunisian yearly crops. True of False

3. What is the population of Libya?

THE NETWORK DATABASE

This database discusses various issues pertaining to the countries of North Africa. The database contains roughly 187 screens which have been connected in a network configuration. That is, a link was established between any two screens where the material in one screen refers to information in another. The links were created based on a "keyword search" strategy; every word in each screen was scanned to see if another screen existed with a title corresponding to a given word. For instance, if you are currently at a screen discussing the climate of the Western Sahara, you may see five linked terms. These links take you to other screens which may have no explicit relationship with the climate of Western Sahara, but are connected since the keyword appeared in the screen discussing the climate of Western Sahara. Please note that these links may or may not be in context. The links are embedded within the text and are clearly marked with a dark rectangle (as you see on the screen). Clicking inside the rectangle will take you to the referent screen. No further explanation shall be given regarding the various linkages within the system; as you perform the tasks you will become familiar with the database.

In performing the tasks, your strategy should be to click on the appropriate link indication to access the screen which you think may contain the answer to the query. If the answer is not found you should move about the database, using the available links, to reach the screen(s) that might provide the answer. The answer to some queries will require information from only one screen, while others may require information from two screens.

The buttons in the lower left-hand corner of the screen are as follows:

The **Prev. Screen** button will take you to the last screen you were at.

The **Path History** button will regressively access each screen you have been to. That is, it will move backward one screen-at-a-time. So if you click on this button once you will be taken to the last screen you were at. If you click again you will be taken to your location from 2 screens past, and so on for up to 42 screens.

The **Glossary** button accesses a glossary of link terms. Each of the available link titles (the titles in the buttons) are listed in alphabetical order and a description of the information contained in the referent screen is included.

The **Home** button will take you directly to the top screen in the database (i.e., the North Africa screen).

The **Answer** button is used for indicating the answer. When you believe you've obtained the answer click on this button. An answer screen will appear; simply position the cursor in the appropriate field and type in the answer.

The experimenter will go through the system with you; take 5 minutes to "play around" in the system and familiarize yourself with it. Try clicking on link indications to see where they lead.

Please attempt the following practice queries:

1. The largest country in North Africa exports 5 agricultural products. True or False

2. Wheat and barley constitute as much as 97% of Tunisian yearly crops. True or False

3. What is the population of Libya?

THE COMBINATION DATABASE

The database contains roughly 187 screens which have been connected in a combined hierarchical and network configuration; it discusses various issues pertaining to the countries of North Africa. In the hierarchical configuration, five major categories exist within each country: demographics, economy, education, geography, and government. No further explanation shall be given regarding the various categories within the hierarchy; as you perform the tasks you will become familiar with the database. For the network connections, a link was established between any two screens where the material in one screen refers to information in another; these links are not dependent on the hierarchy at all. A link can go from any screen (at any location) to any other screen. The links were created based on a "keyword search" strategy; every word in each screen was scanned to see if keywords existed, then a link was created between the current screen and the screen with a title corresponding to the keyword. For instance, if you are currently at a screen discussing the climate of the Western Sahara, you may see five linked terms. These links take you to other screens which may have no explicit relationship with the climate of Western Sahara, but are connected since the keywords appeared in the screen discussing the climate of Western Sahara. Please note that these links may or may not be in context. The links are embedded within the text and are clearly marked with a dark rectangle (as you see on the screen). Clicking inside the rectangle will take you to the referent screen.

In performing the tasks, your strategy should be to click on the appropriate link indication (located either in the upper left-hand corner of the screen or within the text itself) to access the screen which you think may contain the answer to the query. If the answer is not found you should move about the database, using the available links, to reach the screen(s) that might provide the answer. The answer to some queries will require information from only one screen, while others may require information from two screens.

The buttons in the lower left-hand corner of the screen are as follows:

The **Prev. level** button will take you to the previous level in the hierarchy.

The **Path History** button will regressively access each screen you have been to. That is, it will move backward one screen-at-a-time. So if you click on this button once you will be taken to the last screen you were at. If you click again you will be taken to your location from 2 screens past, and so on for up to 42 screens. It does not matter whether you used hierarchical or network links.

The **Glossary** button accesses a glossary of link terms. Each of the available link titles (the titles in the buttons) are listed in alphabetical order and a description of the information contained in the referent screen is included.

The **Home** button will take you directly to the top screen in the database (i.e., the North Africa screen).

The **Answer** button is used for indicating the answer. When you believe you've obtained the answer click on this button. An answer screen will appear; simply position the cursor in the appropriate field and type in the answer.

The experimenter will go through the system with you; take 5 minutes to "play around" in the system and familiarize yourself with it. Try clicking on link indications to see where they lead.

Please attempt the following practice queries:

1. The largest country in North Africa exports 5 agricultural products. True or False

2. Wheat and barley constitute as much as 97% of Tunisian yearly crops. True of False

3. What is the population of Libya?

THE LINEAR DATABASE

This database contains roughly 160 screens discussing various issues pertaining to the countries of North Africa. The six countries are listed in the upper left-hand corner of the "North Africa" screen. Each country has its own section in the database, in which you will find five major categories: demographics, economy, education, geography, and government. Each of these sections has 5 subsections and all are organized as in a book. That is, demographics is the first page, followed by five pages of more detailed subsections, then comes the economy screen, and so on. No further explanation shall be given regarding the various categories within the system; as you perform the tasks you will become familiar with the database.

In performing the tasks, your strategy should be to get to the appropriate screen(s) which you think may contain the answer to the query. If the answer is not found you should move about the database to reach the screen(s) that might provide the answer. The answer to some queries will require information from only one screen, while others may require information from two screens.

The buttons in the lower left-hand corner of the screen are as follows:

The **Contents** button will provide you with a table of contents of the available screens.

The **Path History** button will regressively access each screen you have been to. That is, it will move backward one screen-at-a-time. So if you click on this button once you will be taken to the last screen you were at. If you click again you will be taken to your location from 2 screens past, and so on for up to 42 screens.

The **Glossary** button accesses a glossary of screens. Each of the screen titles are listed in alphabetical order and a description of the information contained in the screen is included.

The **Home** button will take you directly to the top screen in the database (i.e., the North Africa screen).

The **Answer** button is used for indicating the answer. When you believe you've obtained the answer click on this button. An answer screen will appear; simply position the cursor in the appropriate field and type in the answer.

The arrow buttons on the top are used for paging back and forth one screen at-a-time and the arrow buttons on the bottom are used to page back and forth one subsection at-a-time.

The experimenter will go through the system with you; take 5 minutes to "play around" in the system and familiarize yourself with it. Try clicking on link indications to see where they lead.

Please attempt the following practice queries:

1. The largest country in North Africa exports 5 agricultural products. True or False

2. Wheat and barley constitute as much as 97% of Tunisian yearly crops. True of False

3. What is the population of Libya?

Appendix V. Modified QUIS.

FINAL QUESTIONNAIRE

Please rate the system with regard to the following questionnaire items. Remember that the system is under study, not you; your ability to use the system is heavily influence by system characteristics. So be frank in your ratings. Please mark all answers on the OpScan sheet.

OVERALL REACTIONS TO THE SOFTWARE

1. Overall the system was
terrible wonderful
1 2 3 4 5 6 7 8 9

2. Overall the system was
difficult easy
1 2 3 4 5 6 7 8 9

3. Overall the system was
frustrating satisfying
1 2 3 4 5 6 7 8 9

4. Overall the system offered
inadequate power adequate power
1 2 3 4 5 6 7 8 9

5. Overall the system was
dull stimulating
1 2 3 4 5 6 7 8 9

6. Overall the system was
rigid flexible
1 2 3 4 5 6 7 8 9

SCREEN

7. Characters on the computer screen are
hard to read easy to read
1 2 3 4 5 6 7 8 9

8. Highlighting on the screen simplifies task
not at all very much
1 2 3 4 5 6 7 8 9

9. Organization of information on screen is
confusing clear
1 2 3 4 5 6 7 8 9

10. Sequence of screens is
confusing clear
1 2 3 4 5 6 7 8 9

TERMINOLOGY AND SYSTEM INFORMATION

11. Use of terms throughout the system was
inconsistent consistent
1 2 3 4 5 6 7 8 9

12. Computer terminology is related to the tasks you were doing
never always
1 2 3 4 5 6 7 8 9

13. Position of messages on screen
consistent inconsistent
1 2 3 4 5 6 7 8 9

14. Computer keeps you informed about what it is doing
never always
1 2 3 4 5 6 7 8 9

15. Error messages were
helpful unhelpful
1 2 3 4 5 6 7 8 9

LEARNING

16. Learning to operate the system was
difficult easy
1 2 3 4 5 6 7 8 9

17. Exploring new features by trial and error
difficult easy
1 2 3 4 5 6 7 8 9

18. Remembering names (if you had to) and use of commands
difficult easy
1 2 3 4 5 6 7 8 9

19. Tasks can be performed in a straight-forward manner
never always
1 2 3 4 5 6 7 8 9

20. Help messages on the screen were
unhelpful helpful
1 2 3 4 5 6 7 8 9

21. Supplemental reference materials were
unhelpful helpful
1 2 3 4 5 6 7 8 9

SYSTEM CAPABILITIES

22. System speed was
too slow fast enough
1 2 3 4 5 6 7 8 9

23. System reliability
unreliable reliable
1 2 3 4 5 6 7 8 9

24. System tended to be
noisy quiet
1 2 3 4 5 6 7 8 9

25. Correcting your mistakes (if any) was
difficult easy
1 2 3 4 5 6 7 8 9

26. Experienced and inexperienced users' needs are taken into consideration
never always
1 2 3 4 5 6 7 8 9

TASK PERFORMANCE

27. In general, performing the tasks with this system was
difficult easy
1 2 3 4 5 6 7 8 9

28. Getting to the desired screens was
confusing clear
1 2 3 4 5 6 7 8 9

29. During task performance you
lost your place never lost place
1 2 3 4 5 6 7 8 9

30. Please provide an overall rating (of 1 to 9, with 1 being high and 9 being low) of the system for performing the types of tasks you performed today. Indicate your rating on the opscan sheet.

Appendix VI. Post-hoc Test for the TCT Task Type x Required Links Interaction.

TABLE VI-A. Newman-Keuls Post Hoc Test for TCT Task Type x Required Links Interaction ($p < 0.05$) in Experiment A. TCTs with identical letters are not significantly different.

REQUIRED LINKS	TASK TYPE	TCT (s)	
2-Link	Hierarchical	66.26	A
2-Link	Relational	78.33	A
4-Link	Hierarchical	86.86	A
6-Link	Relational	163.97	B
4-Link	Relational	164.52	B
6-Link	Hierarchical	237.60	C

Appendix VII. QUIS MANOVA Results.

TABLE VII-A. Obtained MANOVA parameters for each item of the QUIS.

QUESTIONNAIRE ITEM	System MS	Error MS	F	p
1. Overall the system was terrible--wonderful	12.396	1.454	8.53	0.0001
2. Overall the system was difficult--easy	3.229	1.763	1.83	0.1511
3. Overall the system was frustrating--satisfyir	11.724	2.839	4.13	0.0099
4. System has inadequate power--adequate power	24.354	1.581	15.40	0.0001
5. Overall the system was dull--stimulating	10.938	3.461	3.16	0.0311
6. Overall the system was rigid--flexible	11.766	2.932	4.01	0.0114
7. Characters hard to read--easy to read	0.682	3.295	0.21	0.8911
8. Highlighting simplifies task none--very much	Subjects provided no rating			
9. Organization of information confusing--clear	1.188	2.191	0.54	0.6552
10. Sequence of screens is confusing--clear	1.391	2.441	0.57	0.6371
11. Term usage was inconsistent--consistent	3.807	0.824	4.62	0.0057
12. Terminology is never--always related to task	6.807	7.703	0.88	0.4548
13. Position of messages is inconsistent--consist	6.849	11.174	0.61	0.6093
14. Computer keeps you informed	4.229	4.979	0.85	0.4724
15. Error messages were unhelpful--helpful	2.432	9.836	0.25	0.8631
16. Learning to operate was difficult--easy	2.391	1.395	1.71	0.1737
17. Exploring new features was difficult--easy	3.792	2.094	1.81	0.1548
18. Recall and use of commands was difficult--eas	5.229	2.538	2.06	0.1149

TABLE VII-A (Cont.).

QUESTIONNAIRE ITEM	System MS	Error MS	F	p
19. Tasks never--always done straight-fowardly	2.729	2.121	1.29	0.2871
20. Help messages were unhelpful--helpful	4.224	6.403	0.66	0.5801
21. Reference materials were unhelpful--helpful	11.292	8.198	1.38	0.2584
22. System speed was too slow--fast enough	42.599	2.819	15.11	0.0001
23. System was unreliable--reliable	1.432	2.922	0.49	0.6904
24. System tended to be noisy--quiet	0.604	1.219	0.50	0.6866
25. Recovering from errors was difficult--easy	5.433	2.464	2.20	0.0967
26. Expert/novice needs never--always considered	1.792	2.773	0.65	0.5884
27. Perfroming tasks was difficult--easy	5.932	1.868	3.18	0.0304
28. Getting to desired screen was confusing--clea	6.354	2.148	2.96	0.0394
29. You lost your place--never lost place	5.099	3.153	1.62	0.1948
30. Provide overall rating from 1 (low) to 9 (hic	41.833	0.937	44.65	0.0001

VITA

Michael F. Mohageg was born on September 30, 1964, in Burbank, California. He received a Bachelor of Arts degree in Psychology from the California State University at Northridge (1986) and a Master of Science degree in Industrial Engineering and Operations Research from Virginia Polytechnic Institute and State University (1989).

Mr. Mohageg has been actively involved in HCI as both a student and a researcher for the past 4.5 years. His general areas of interest include hypertext/hypermedia, graphical user interfaces, computer graphics, and input devices. He is a member of Alpha Pi Mu and a student member of the Human Factors Society (HFS), and has served as treasurer of the VPI & SU student chapter of HFS (1988). Mr. Mohageg will continue his human factors career with the Science and Technology Division of the NYNEX Corporation in White Plains, New York.

Signature