READERS' COMPREHENSION AND STRATEGIES IN LINEAR TEXT AND HYPERTEXT

Peter W. Foltz

Institute of Cognitive Science
Box 344
Univeristy of Colorado
Boulder, CO 80309

ICS Technical report #93-01

ABSTRACT

Hypertexts present methods to read online texts that are different from those available when reading standard linear texts. Hypertexts give readers more flexibility in choosing paths through the text and in finding relevant information. However, research in hypertext has often shown little or no advantages over the equivalent linear text.

This research applied text comprehension theory to make comparisons of the readers' comprehension and strategies in linear text and hypertext. The theory hypothesized that hypertext would break down the coherence of the text, adversely affecting the readers' comprehension.

To test this hypothesis, two experiments were run. In the first experiment, subjects read a textbook chapter in one of three text formats: linear text, hypertext, or coherent hypertext. Subjects reading goals were manipulated so that half looked for specific information while the other half read for general knowledge. Although the reading goals affected what subjects looked for in the text, comprehension was equivalent for the three text formats. An analysis of the subjects' reading strategies indicated that despite the different text formats, subjects used a similar reading strategy. Subjects tried to maintain a coherent path through the text, which resulted in subjects seeing a similar coherent representation of the text.

The second experiment examined the strategies used in the hypertext format. Verbal protocols of the subjects indicated that they used reading heuristics and signals in the text to guide them through the text in a coherent manner. Even when subjects jumped to different places to find specific information, they still made these jumps coherently.

The results of both experiments provide evidence for the dominance of maintaining global coherence in a text. Subjects presented with opportunities of reading a hypertext non-coherently, nevertheless chose to use strategies to maintain a coherent macrostructure of the text. These results suggest areas of improvement for models of hypertext readers' comprehension and strategies and implications for the design of successful online text systems.

CONTENTS

TABLES

Table

FIGURES

Figure

CHAPTER I

INTRODUCTION

> You shall see them on a beautiful quarto page, where a neat rivulet of text shall meander through a meadow of a margin.
> *Richard Brinsley Sheridan, The School for Scandal (1777)*

Hypertext presents a way to read online text that differs from reading standard linear text. Text is typically presented in a linear form, in which there is a single way to progress through the text, starting at the beginning and reading to the end. However, in hypertext, information is represented in more of a semantic network in which multiple related sections of the text are connected to each other. A user may then browse through the sections of the text, jumping from one text section to another. This permits a reader to choose a path through the text that will be most relevant to his or her interests.

The concept of using these associative paths to retrieve and read information has caused great excitement. The promise of universally available hypertexts has been touted as "a seamless and reunited computer world" (Xanadu advertisement), having "the potential to become a significant application area; equaling or perhaps exceeding that of word processing, spreadsheets and general database applications" (Begoray, 1990, p. 121), and "because hypertext has the power to change the way we understand and experience texts, it offers radical promises and challenges to students, teachers and theorists of literature" (Landow, 1989, p. 174). Using associative retrieval paths is similar to the way retrieval is performed from human memory, and this may be part of the appeal to hypertext researchers and developers when they state that hypertext systems will improve a user's ability to find and use information.

While a variety of hypertexts have been developed over the past 20 years, it is often not clear whether there are strong advantages for hypertext. Research in hypertext has sometimes failed to show a significant advantage for reading a hypertext compared to the equivalent text in linear form. In addition, the effectiveness of various features that can be used in hypertexts can vary greatly depending on the domain and content of the text and the goals of the reader. Up to this point, no standards or definitive rules exist on how to develop an effective hypertext. However, because hypertext encompasses such domains as user-interface design, psychology, education,

and information retrieval, theory from these domains can be applied to hypertext in order to aid in the understanding and development of effective hypertexts.

The research in this thesis approaches the evaluation of hypertext from the perspective of text comprehension. There has been a large body of research developed over the last 40 years studying linear text comprehension, both from a theoretical and an evaluative standpoint. Text research has permitted predictions of comprehension based on such factors as the structure of the text, the background knowledge of the reader, and the reader's abilities. Thus, the goal of this research is to evaluate hypertext using the guidelines and methods from text comprehension research.

Although hypertext can be defined in many different ways, ranging from information retrieval systems to adventure games, this research looks at several specific factors of hypertext. In this way, it is not as much an evaluation of hypertexts in general or of the issues of information seeking in hypertext as it is of the role of reading and comprehension in hypertext. This evaluation permits answers to several questions crucial to the development of online text systems: what are some of the constraints on the design of texts in hypertext form, what are the strategies used by readers of hypertexts, and why has research shown very few advantages of hypertext over linear text.

CHAPTER II

HYPERTEXT: TEXT MEETS COMPUTER TECHNOLOGY

> The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of the square-rigged ships.
> *Vanevar Bush, As we may think (1945)*

Writing, which dates back to around 6000 BC, has long been the primary method for both communicating and storing information. While books and libraries provide people with moderately easy access to information, computers have made the retrieval of information more efficient. This improvement is largely due to the ability to have information linked with other information or indexed in an easily retrievable manner. Computers can automate much of the indexing that had originally been performed manually, and thus permit people to search for information using these indexes. These retrieval systems have been used primarily for card catalog and other indexed retrieval systems. Automatic indexing and retrieval also makes possible the encoding of links between related pieces of information and enables people to retrieve information that is related to what they are currently reading. This is the primary feature of hypertext.

Although hypertexts represent the computerized version of encoded links between related pieces of information, non-computerized hypertexts have been created through the centuries. The Talmud, a compilation of Jewish oral law from the 5th century AD, contained many hypertext structures. The Talmud was written on large pages with small amounts of the basic text written in the center. Around the basic text were written five to ten other texts made up of commentaries on the original text, commentaries on the commentaries, notes, and references to the Bible. The cross references from the texts to the relevant commentaries were signaled in the texts through an elaborate set of symbols (icons). Thus, a reader of the original text could use the icons to look up information in related commentaries on the same page, or know where to look in the bible to find the related information. Remarkably, the creation of the Talmud was entirely crafted by hand, and all cross-referencing that a reader must do would be done manually. Today, hypertexts on computers allows readers to follow such links automatically. Thus, hypertext removes some of the constraints of either looking up

information in an index and then going to find it, or as in the Talmud, requiring large pages with all the relevant information on it with an elaborate cross-referencing system.

History and Features of Hypertext

The earliest examination of automated hypertext concepts were by Vanevar Bush in the 1930s and 1940s (Bush, 1933, 1945). Bush identified a problem for researchers who want to search and put together disparate pieces of published research. During this period, scientific literature was expanding at a great rate and it was becoming more difficult to find the relevant information (Nyce & Kahn, 1991). As a solution to this problem, he envisioned a machine called *Memex* that could be developed and used with the goal of augmenting human intellect.

With a Memex, a user could file all information acquired during the day, such as the information from books, records and communications into the machine in microfilm form. This information would be stored using index codes and could then be rapidly retrieved through using the same index codes. For the retrieval, Bush envisioned a mechanical analogy to the human brain.

> When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. ... The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain. (Bush, 1945, p. 106)

> [A]ssociative indexing, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. The process of tying two items together is the important thing. (Bush, 1945, p. 108)

Bush devised Memex as a desk containing two screens for viewing microfilm documents, a keyboard for entering index codes and a set of levers for moving through documents. When two separate files were brought up on the screens, they could be associated or linked with each other through a single keypress. Then, the next time one of the documents was retrieved, it would be simple to bring up the other. In this manner, a researcher could build a trail through a series of documents while researching a topic and then later bring up that same trail if he/she needed to refer to the information again.

Bush also saw that these trails could be literary constructs in themselves. People with the same libraries could exchange trails so other people could follow the same trails in order to learn the same information. Thus, it was

not only the information that was useful, but the order and structure of the presentation of that information. He hypothesized that this would be useful for such people as physicians and patent attorneys who could follow the trails of arguments in documents in order to study previous cases. Due to the continued need of these trails, he saw that a new career could be developed of "trailblazers", whose job would be to develop trails through relevant documents, and then sell the trails with the documents.

Bush's proposed system was still far ahead of the technology of the day, and the implementation of a hypertext would have to wait until computer technology existed. Almost 20 years after Bush's vision of Memex, the first hypertext system was actually implemented on a computer system (Engelbart, 1963; Engelbart & English, 1968). Since then, a number of hypertext systems have now been developed for a variety of domains and tasks.

Hypertext takes a wide range of forms and has been implemented in many different ways. Nelson (1967), one of the early pioneers of hypertext, has defined it as "a combination of natural language text with the computer's capacity for interactive branching or dynamic display... of nonlinear text... which cannot be printed conveniently on a conventional page." From this quote, we see some of the basic defining features of hypertext that differentiate it from linear text and also information retrieval systems. The primary feature of hypertext is that the text information is no longer in a linear form and thus is non-sequential. In linear text, there is a single fixed order to reading the text that has been set down by the author. In hypertext, there may be multiple orders to reading the text, with the reader choosing the path that is best for the reader's needs. The ability to read a text non-linearly makes intuitive sense, since we often do not read large amounts of information in books linearly; we may read a section, look something up in the glossary or index or refer back to earlier sections. Many existing texts are likewise not designed for linear reading, such as encyclopedias and technical manuals.

A second defining feature of a hypertext is the associative connections between related items. These connections, either created by the author or by the computer, allow a reader to move from one topic to a related topic. In this manner, reading is like trains of thought, in which information items appear in the proper context. A reader then reads by moving along the desired connections.

Because there are multiple connections through the text, a third feature of a hypertext is its dynamic representation. All readers of the same text may not see the same things. The text may be automatically represented differently to different people based on their background knowledge for their task.

A final feature in many hypertext systems is the ability to do some form of search in the text. While books have hand crafted indexes, a hypertext can have a search mechanism which will take a reader to the parts of a text in

which a particular word or phrase occurs. This permits the reader to find the relevant information in a text in a fast and efficient manner.

Although the above features are not all inclusive or exclusive of hypertext systems, they represent the primary features used in most hypertexts. These features therefore supply much flexibility to the reader when compared to reading linear, text such as books. Clearly some of this flexibility does exist in books (e.g. table of contents and indexes), but it is not as widely used or exploited. Hypertext permits readers to use these features automatically rather than requiring readers to manually refer to them as needed. This additional flexibility gives much more control to the reader in determining the order that the text is to be read, and allows the reader to read the text as if it were specifically tailored to the reader's background and interests. This flexibility does promise this advantage of personalization and eases the burden of finding information, however a question that remains is, whether this flexibility is actually good or useful to the reader.

Research in Hypertext

The development of hypertext systems has created research into how to design and use hypertexts. There have been several major areas of research which include: the development of the underlying representations of information and connections (e.g., Furuta & Stotts, 1989; Botafogo & Shneiderman, 1991), methods of connecting, structuring and retrieving the information (e.g., Croft & Turtle, 1989; Crouch, Crouch, & Andreas, 1989), designing hypertexts for supporting argumentation (e.g., Conklin & Begeman, 1989; Fischer, McCall, & Morch, 1989), studying information seeking in hypertexts (e.g., Marchionini & Shneiderman, 1987; Weyer, 1982), and the role of rhetoric and writing in hypertexts (e.g., Bolter, 1991; Britton & Glynn, 1989).

While there have been a variety of areas of research on hypertext, one of the primary goals of hypertext research is to evaluate hypertexts and hypertext features in order to bring about improvements. This thesis will focus on two basic areas of research that examines issues of the evaluation of hypertexts. The first area of research has been to compare hypertext to linear text in different situations. The second area has been to compare a variety of hypertext features against each other in order to determine what features will be the most effective and in what type of situations. From these two types of research, we can then get an idea how well hypertexts have succeeded and what hypertext features are the most effective.

Comparisons of Linear Text and Hypertext

Research comparing hypertext to linear text provides some measure as to what situations will better suit the presentation of textual information with a

hypertext. While there have been claims of the superiority of hypertext over linear text (e.g., Martin, 1990; Nelson, 1967), research into this area has not been uniformly successful in showing an advantage for hypertext. Gordon, Gustavel, Moore, & Hankey (1988) presented subjects with magazine articles that were either in their original linear form, or converted to a hierarchical hypertext using the text structure to form the hierarchy. They found overall that subjects had a better free recall of information when reading the linear text. When reading general interest articles, subjects had a better understanding of the macrostructure of the text when reading it in linear format than hypertext, however there were no differences between linear and hypertext format for reading technical articles. Shneiderman (1987a) found that subjects using the Hyperties hypertext version of a set of historical articles took about the same time to find answers to questions, although subjects were faster with the linear text version if the answer could be found near the start of the article.

These studies show some advantage for the linear text over the hypertext, but several other studies using the SuperBook hypertext system have shown advantages for hypertext. Egan, Remde, Gomez, Landauer, Eberhardt & Lochbaum (1989) found that with their SuperBook system, subjects were faster at finding question answers, answered more questions correctly, and wrote better essays than with the same text presented in linear form.

Overall, the results of the studies comparing linear text to hypertext vary greatly. A large part of the variability in these results may be due to the different tasks, hypertext systems and texts used. Thus, the results indicate that there are several factors that need to be considered in order to determine whether a hypertext or linear text would work best.

Evaluation of Hypertext Features

Because hypertexts have incorporated many new features that are not found in linear text, such as linking, the ability to search for information, guided tours and overview maps, it is often not clear which of the features may be effective in improving the users' ability to read and comprehend the text. Without information about what features work best, it is difficult to know what features to incorporate into the design of a new hypertext. For this reason, there have been studies comparing the various hypertext features.

One factor that may affect reading hypertext is the resolution of the screen. Work by Gould (Gould & Grischkowsky, 1984; Gould, Alfaro, Fonn, Haupt, Minuto, & Salaun, 1987) found that reading time is significantly slowed when reading a text on a computer screen compared to reading the same text on paper. However, when a high resolution screen with anti-

aliased fonts are used, the reading speeds are equivalent. Thus, studies comparing linear text on paper to hypertexts on screens may find some reading differences just due to the resolution of the text rather than any factors of the actual structure of the text. For this reason, research that makes comparisons of hypertext and linear texts should ensure that they both have equivalent resolution of the text.

With the multiple paths that are possible when reading a hypertext, there is a greater navigation load on the reader than with linear text. In order to simplify the reader's task then, additional navigation features must be provided. In a linear text, a table of contents and an index are the two typical navigational features. Since linear text is often organized hierarchically, the table of contents serves as an outline with pointers to pages for each entry. However, in a hypertext, the structure may be much more complicated. For that reason, a common feature in a hypertext is a map. A map permits a reader to see a representation of the text's structure along with the relationships between the different text sections. Monk, Walsh & Dix (1988) found that with a map, subjects were much faster at being able to answer questions in a hypertext than with no map. However, in a similar study, Hammond & Allinson (1989) found no differences in answer accuracy or task time with or without a map in a hypertext. Nevertheless, they did find that with a map subjects did tend to visit more of the text nodes than the subjects who did not have a map. There were large differences in the style of navigation based on the subject's task. Subjects who were instructed to just explore the information space tended to use a guided tour facility, while subjects with the more directed tasks of answering particular questions used mostly the map and index to help navigate.

Thus, we see that there are varying results as to whether certain features aid the user of a hypertext . In addition, factors such as the user's goals and background knowledge can influence whether a particular feature is useful. Nielsen (1989) performed a meta-analysis of 92 benchmark measures that had been taken from usability issues that had been tested in hypertext research. In all of the measures he examined, he found that there were actually very few results which had large effects. From this analysis, he concluded that there are actually very few studies that have shown the real world impact of hypertext systems. In addition, the lack of strong effects in this area of research could mostly be explained due to large individual differences among users, tasks and texts. Thus, while there has been a lot of research in hypertext, few concrete results exist to provide strong evidence on how well hypertexts will work and guidelines on how to create effective hypertexts.

Theory-Based Approaches to Hypertext Design

The studies described previously shed some light on what features may be used to develop a good hypertext system, however, they are very insular, only examining particular features with particular texts. Without an ability to generalize outside of the texts and tasks, with every new feature or text developed, another evaluation study would need to be performed to determine its effectiveness. An alternative approach is to use a theoretical background to drive the design of the hypertext. This then would permit comparisons of features through using the theory to make the performance predictions. Theory-based approaches have been widely used in human-computer interaction (e.g. Card, Moran & Newell, 1984; Kieras & Polson, 1985; Polson & Lewis, 1990). These approaches use a theoretical model of the user and system in order to derive predictions of usability. They typically involve modeling such factors as the users' knowledge, possible states of the computer, and possible actions a user can take, and have had some success at predicting usability of systems and of particular features. This suggests that theory-based design can similarly be done for hypertext systems.

There have been a few theoretical approaches to studying and designing hypertext. While not based on some of the strict modeling criteria used in the modeling mentioned above, these approaches do use theoretical bases or guidelines for their development and testing. One approach has been to use user interface guidelines, such as dialog rules, to determine how a hypertext interface should be designed (Shneiderman, 1987b; Hardman & Sharrat, 1989). This approach permits pieces of the interface of a hypertext to be developed based on proven guidelines. A second approach has been to use "formative design evaluation" (Egan, Remde, Gomez, Landauer, Eberhardt & Lochbaum, 1989, Landauer, Egan, Remde, Lesk, Lochbaum, & Ketchum, 1992). In formative design evaluation, development and changes of features are based on psychological guidelines. Behavioral studies are then performed on the system in order to determine what features can be improved and in what ways. This iterative design method allows a comparison of the features and quantitative measures of the improvement from previous versions. A third approach has been to develop hypertexts from the point of view of information retrieval theory. In this approach, models from information retrieval have been applied in order to determine how to structure the information. These models use methods such as, probabilistic models of retrieval (Croft & Turtle, 1989), hierarchical clustering (Crouch, Crouch & Andreas, 1989), and petri nets (Furuta & Stotts, 1989) in order to connect the nodes of information.

While these approaches have helped improve the design of hypertexts, one area that has been neglected is to consider the text in terms of its discourse cues. Charney (1987) has suggested this area as a possible area for improving hypertexts. However, thus far, there has been little theoretical research in this area. This may be partially due to the fact that up to this point

in the history of hypertext, developers of hypertexts tend to be primarily more computer scientists than researchers with skills and background in issues of text comprehension. Thus, the research described below will apply a new theoretical approach to hypertext, examining it in terms of text comprehension and discourse cues.

Guidelines for Developing Hypertexts

Another source of how to develop hypertexts comes from development guidelines. Several books and articles have been published containing guidelines and rules for hypertext development. These guidelines concentrate on issues such as how much text should be contained in a node, what hypertext features to use, and how the information should be structured. However, most of these guidelines do not provide concrete design rules based on theory, but instead present abstract rules based more on common sense. (e.g., "For the short term, the best recommendation probably is to pay close attention to the authoring principles implicit in other writers' hypertext and try to emulate the principles you like." (Nielsen, 1990, p. 164)). The fact that the rules are fairly abstract is not surprising given that writing a hypertext is as difficult, if not more difficult, than writing linear text. With linear text, we do have some guidelines, but they still are somewhat abstract, and effective writing depends more on large amounts of practice than just being able to follow a set of writing rules. In addition, the style of writing will depend greatly on the points the author wants to express and the domain in which the author is writing.

The existing hypertext guidelines recommend dividing the text into small sections and then creating a clear set of links between these sections.

> The reader of a good hyperdocuments finds it useful to be able to follow links directly to subjects that interest him... Success in authoring Hyperdocuments then depends on dividing the subject matter into self-contained fragments... If removed from its document, a nugget can usually be understood in its own right" (Martin, 1990, p. 49).

Most guidelines further recommend that the structure be organized into a hierarchical form (e.g., Akscyn, McCracken & Yoder, 1988; Martin, 1990). This hierarchical representation is argued to be similar to the reader's mental model of the text and thus should help them in seeing a coherent organization of the text. From this hierarchical organization, cross-hierarchical links can then be added to permit readers to move between related items that are not directly hierarchically related.

CHAPTER III

TEXT COMPREHENSION RESEARCH

> Of all the need a book has, the chief need is that it be readable
> *Anthony Trollope, An Autobiography (1883)*

Over the past 30 years, there has been a large amount of research in text comprehension, primarily in the fields of psychology and education. The goals of text comprehension research are to understand what factors of the reader and the text influence the ease of comprehending a text and to make some predictions on how easy a text will be to comprehend. Through modeling of the text and the reader's knowledge and abilities to make these predictions, researchers have been able to both develop better texts and a better understanding of the human comprehension processes.

Research in text comprehension has examined a variety of factors that influence comprehension. These factors include: the role of coherence in a text, the role of the readers' background knowledge, the role of the narrative schema of the text, and the role of the reader's cognitive abilities. In addition, studies have examined the strategies readers use when going through a text and the role these strategies can play for comprehending a text.

This chapter will review research performed that describes some of the factors that influence comprehension and how these factors play a role in the comprehension of text. The factors described in this chapter are not inclusive of all major factors that influence text comprehension. However, the factors chosen were expected to play an equally important role in the comprehending of hypertexts and linear texts.

Predictions of Comprehension

Modeling Comprehension

Researchers in the field of text comprehension have used user models to predict what information will be learned from a linear text. This dissertation will examine the comprehension and goals of readers of a hypertext and linear text using the Kintsch model of text comprehension as its basis (Kintsch, 1988; Van Dijk & Kintsch, 1983). The Kintsch model has been used

for predicting the comprehension of text based on such factors as what features will be remembered from the text (Van Dijk & Kintsch, 1983), the role of background knowledge (Britton & Gulgoz, 1991; van Dijk & Kintsch, 1983), the role of coherence and readability (Kintsch & Vipond, 1979), and also goal planning in such domains as computer mail systems (Mannes & Kintsch,1991) and the Unix operating system (Doane, Kintsch & Polson, 1990).

When reading a text, processing occurs at many levels. These levels range from the low level processes of recognizing individual words up to high level processes of deriving the gist of the information in the text. These processes work together simultaneously to extract meaning from the text. Meaning however, is represented at different levels.

In the Kintsch model, the reader's memory for text is represented at 3 levels, a surface representation of the words and sentences, the meaning of the text (textbase) and a general representation of what is described by the text incorporating outside background knowledge (situation model). As text is read, the text is incorporated into the readers representation of the information. Information from the surface representation of the text is quickly lost (e.g., Bransford & Franks, 1971). However, some of the abstracted information from the surface structure of the text is incorporated into the textbase, represented as propositions. Propositions serve as semantic primitives representing the information acquired (e.g. Kintsch, 1974; Fodor, Fodor & Garrett, 1975).

The propositions are connected to each other in the textbase through coherence relations. Coherence in a text may be represented in many forms, including; *syntactic*, in which the use of the language expresses coherence through the use of such things as pronouns, *stylistic*, in which the discourse makes use of a similar style and lexical choices, and *pragmatic*, in which the choice of words can be used for a particular pragmatic context. In this research, semantic coherence is examined. In semantic coherence, constituents of the text will be coherent if they share some form of semantic relatedness in the discourse. The amount of coherence is therefore represented by the amount of shared meaning and referential relations. These coherence relations are based on standard rhetorical devices in the text, such as causality, use of pronouns, and word repetition. The semantic coherence can be represented both at local level of individual propositions in the textbase and at the global level of the macrostructure of the text.

The mental representation of these interconnected propositions in the textbase (coherence graph) takes the form of a hierarchical structure with higher level concepts represented as super-ordinate propositions which are connected to lower level concepts represented as subordinate propositions (Kintsch & van Dijk, 1978). Recall of propositions follows this hierarchy, with propositions from the upper part of the tree being more likely to be recalled than lower level propositions (Britton, Meyer, Hodge & Glynn, 1980; Kintsch & Keenan, 1973; Meyer, 1973). At a higher level, readers generate a

macrostructure or gist of the text. The macrostructure is the result of the readers' inferential processes, with readers forming hypotheses of the overall meaning of paragraphs, chapters and whole books. In this manner, the macrostructure incorporates the reader's background knowledge with the text in forming these hypotheses. The resulting macrostructure of the text is similar to the textbase in that it is also represented as a hierarchical structure with higher level concepts represented at the top.

The Role of Coherence

The process of incorporating propositions into the coherence graph is a process of maintaining coherence. Propositions that have overlapping arguments, and thus are semantically related, create coherence. However, if the current proposition being processed does not share arguments with propositions in short-term memory, then a bridging inference must be made by the reader in order to maintain coherence (e.g., Kintsch & van Dijk, 1978). When a reader makes a bridging inference, the reader must use knowledge from his or her situation model in order fill in the missing information. Writers often assume readers will have the appropriate background knowledge to make the proper bridging inferences, and knowledgeable readers can make these inferences automatically. However, if a reader does not have the proper background knowledge, these inferences consume additional resources of the reader, typically resulting in lower comprehension. Thus, the amount of coherence in the text, can be used to make predictions of comprehension (e.g., Kintsch & Vipond, 1979; Miller & Kintsch, 1980).

One application of this predictive power has been to improve the comprehensibility of texts. Britton & Gulgoz (1991) used the Kintsch model to identify locations in a text where a reader would have to make these bridging inferences in order to maintain coherence of the text. The chosen text was on the topic of the air war in Vietnam and so undergraduates had very little background knowledge to make the correct bridging inferences at these incoherent locations. They created a revised version of the text in which they inserted into the text the inferences that would be needed to be made at the various locations. They found that undergraduates reading the revised text had significantly better recall and a better mental representation of the text than those who had read the original version.

The Role of Background Knowledge

As a text is read, there is a large cognitive load on the reader as the reader is decoding the text and incorporating the textual information into his or her knowledge base. The key to incorporating the information into the reader's knowledge base is partly dependent on the amount of background knowledge

of the reader. A reader's background knowledge permits the information to be incorporated into pre-existing knowledge structures. So readers who do not have an adequate amount of background knowledge on the subject of a text will have lower comprehension of the text (Spilich, Vesonder, Chiesi & Voss, 1979; Voss, Vesonder & Spilich, 1980). The background knowledge permits the reader to provide a coherence to the text, permitting better bridging inferences between non-coherent sections and also permitting additional elaborative inferences. In addition, there is evidence that the background knowledge influences the processing of information for developing at the situational model level, but not at the propositional level (Fincher-Kiefer, Post, Greene & Voss, 1988). Thus, additional background knowledge of the text makes it easier for developing relevant macropropositions, resulting in a better representation of the text for the reader.

The Role of Narrative Schema

Similar to the background knowledge, the narrative schema of a text can aid in the comprehension of that text. A knowledgeable reader can use the narrative schema to provide a structure used in which to organize the text (e.g., Black & Bower, 1979; van Dijk & Kintsch, 1984). Greeno and Noreen (1974) found that subjects can process lower level sentences faster if those sentences are preceded by higher level sentences. This is because part of the conceptual structure has already been built by the higher level sentences.

When the expected structure is violated, however, comprehension may be impaired. Poulsen, Kintsch, Kintsch and Premack (1979) found that children had a better recall of the macropropositions of a story when the story was shown in its proper order versus in scrambled order. Although the information presented was the same, the scrambled order violated the children's schema of the story. Without the ability to place the story in a narrative schema, the information was not as well retained. Thus, placing a text in a well known narrative schema can ease the comprehension of that text.

The Role of Readers' Abilities

With the large amount of both high-level and low-level processing that must take place in order to comprehend a text, one can expect differences in comprehension based on the reader's abilities. Skilled readers tend to have better skills at exploiting context cues and other textual constraints. They are able to make better hypotheses about the meaning of words (Perfetti & Roth, 1981), and are more responsive to the rhetorical structure of the text (Meyer, Brandt & Bluth, 1980). Poor readers on the other hand do not have as effective decoding skills and instead compensate through using context-

dependent hypothesis testing. Thus, skilled readers are able to use parallel automatic processing to form better hypotheses about the meaning of the text as they read through it and are not as dependent on the contextual cues of the text. In this manner, if contextual cues are missing or are confusing, then the performance of poor readers will be degraded to an even greater extent.

Readers' Strategies in Linear Text

While much of the results described above are based on tasks in which a reader reads through a text in a single linear order, readers can use various strategies for reading through a linear text. It is not only the reader's abilities and static characteristics of the text that determines the comprehension, readers also rely on structural, syntactic and semantic signals in the text. These signals provide evidence for the macro-relevance of the individual sections of text. A variety of studies have examined the role of different types of these text signals on comprehension, including studies of titles (Bransford & Johnson, 1972), frequency of mention (Perfetti & Goldman, 1974), enumeration (Lorch & Chen, 1986), and initial sentences (Kieras, 1981).

These signals can affect comprehension. They can also determine what a reader looks for in a text and what the reader ignores. A reader may look through a text and find relevant sections based on these signals. and skip over sections that do not seem to be as relevant. In this way, a linear text may not always be read linearly. There has been some research on the strategies used in reading a text. However, reading strategies have not been investigated as much as some of the comprehension processes in reading since there is a lot less control of the conditions and it is harder to determine what information a subject has read and the subjects' motivations for choosing the particular information from the text.

Goldman and Saul (1990) identified a number of strategies used by subjects reading text passages. These strategies were identified both at the global level of the text and at the more local levels. Subjects read individual sentences and could go back and forward through the sentences. At the global level, they identified three approaches, *Once through*, in which subjects read straight through a text, *Review*, in which subjects went to the passage end and then reviewed sentences, and the *Regress* approach, in which they went back to previous sentences throughout the text. They found that subjects almost always used more than one approach to reading a passage. Thus, readers employed a great deal of flexibility in reading the passages. However, their approaches for getting through the text were done in such a manner to establish both local and global discourse coherence. From these results, they proposed a model of strategy competition for reading strategies. In their model, the choices of where to go in the text follow a set of procedural rules. These procedural rules work at both the global and local level to

maintain coherence of the text, to react to the textual features that serve as cues to the coherence relations and to make strategy choices when coherence can not be established.

Thus, reading is not a static process that proceeds in a single order. Readers can employ a variety of strategies for proceeding through a text. These strategies may be determined by a variety of factors, including, the reader's knowledge for the domain, the reader's goals, and the characteristics of the text. However, they still work to maintain the overriding goal of proceeding through the text in a coherent manner.

CHAPTER IV

IMPLICATIONS OF TEXT COMPREHENSION RESEARCH TO HYPERTEXT

> The stumbling way in which even the ablest of the scientists in every generation have had to fight through thickets of erroneous observations, misleading generalizations, inadequate formulations, and unconscious prejudice is rarely appreciated by those who obtain their scientific knowledge from textbooks.
> *James Bryant Conant, Science and Common Sense (1951)*

Predictions of Comprehension

Hypertext differs from linear text in some fundamental ways. A hypertext provides more flexibility to the reader in choosing where to go in the text. A hypertext also provides the reader with more methods to employ in order to find the relevant information in the text and to move through the different sections of the text. However, aside from these differences, the primary goal of both hypertexts and linear texts is to convey textual information in a coherent form to a reader. In this manner, the reader should be able to extract the relevant information from the text that should reflect the intended goals of both the author and the reader. For this reason, we can examine how the results from research in text comprehension can have implications to the design and evaluation of hypertexts.

The Role of Coherence in Hypertext

In a linear text, a writer typically maintains a set of coherent arguments through the text. At the local level, a writer makes words and sentences flow together through common referents. At a global level, a writer similarly makes paragraphs and sections flow from one to the other in a coherent manner. This aids the reader in structuring the information in the text to fit into the knowledge structures of what has been read previously. If there is little global coherence between sections, then the user must make bridging inferences in order to maintain coherence (e.g., Kintsch & Vipond, 1979). For readers without appropriate background knowledge, these inferences can

consume the resources of the reader, typically resulting in lower comprehension.

In a hypertext, at any text section there are usually a variety of other sections to which a reader can jump. However, it may not be possible for a writer to anticipate all the possible places to which a reader may jump and therefore, it may also not be possible to maintain good macrocoherence for all possible links. Although the writer of the hypertext may code all the links, it would still be difficult to write each section so that it would cohere well with every possible section to which a reader could jump. Jumps which are not coherent could result in additional processing load for the reader, as the reader generates the necessary inferences to incorporate the textual information from the new node into what has been previously read. Thus, a many of the possible links in a hypertext may cause difficulty in the reader's comprehension.

One exception to this problem of hypertext coherence may be found in argument-based hypertext systems (e.g., Conklin & Begeman, 1989; Fischer, McCall, & Morch, 1989). These systems take into account the role of coherence through only allowing jumps between nodes in which a coherent argument has been previously set between the two. These coherent links are created through careful hand-coding of all possible links. While argument-based hypertext involves a lot more hand crafting in order to create only these coherent links, it avoids the problems of readers jumping to nodes using links which may cause an incoherent transition.

The Role of Background Knowledge in Hypertext

The amount of background knowledge of the reader can also differentially affect the readers of a hypertext. As discussed previously, readers with background knowledge on the domain of the text will be better at encoding information from the text than those without the background knowledge. The high knowledge readers will have the correct conceptual structures in which to integrate the new information and therefore the reader will tend to have better recall of the text. This will also hold true for readers of a hypertext.

However, due to the flexibility of hypertexts, readers with little background knowledge may have additional difficulties when compared to readers of equivalent linear texts. Readers with background knowledge will already have the correct conceptual structures for the domain. The hypertext structure may therefore be very familiar for them. However, a reader with little knowledge of the domain of the text will not be familiar with the structure of the hypertext. Since one of the concepts of hypertext is to permit more flexibility for the reader in choosing where to go, a low knowledge reader may not be able to accurately choose the relevant text sections. Thus,

low knowledge readers may have additional problems of navigating through the hypertext structure. This problem may not be as evident in linear texts, since the linear text provides a single (default) path to read through the text. Low knowledge readers can always take this path.

The Role of Narrative Schema in Hypertext.

In linear texts, there are a variety of common narrative schemata employed. Most readers are familiar with them and can use their knowledge of a schema in order to help structure and integrate the textual information. However, a narrative schema is primarily dependent on a familiar organization of the presentation of the information. With the novel structures found in hypertexts, much of any familiar narrative schema will not be evident. This can cause difficulties for the readers of a hypertext because they can not effectively organize the textual information into their situation model. While readers of hypertexts may not currently be able to rely on a familiar narrative schema, this may change in the long run. As hypertexts become more accepted and widespread, writers of hypertext may develop standard rhetorical styles. Readers who are then familiar with those rhetorical styles can use that knowledge to help in their structuring of the information in an effective manner.

The Role of Reader's Abilities in Hypertext.

Skilled readers of linear text are more responsive to the rhetorical structure of the text and possess better decoding skills. In contrast, poor readers must rely more on the context to help in decoding. Hypertext may provide fewer context cues than equivalent linear texts. This could be the case in a hypertext in which a reader is not given a lot of information about the structure of the text or about where to go in the text. In a linear text, a poor reader could rely on information from the linear structure for this context, but the same context may not be as evident in the hypertext. On the other hand, a well structured hypertext that provides a map of its structure may provide additional context cues that would not be present in a linear text. In such a case, a poor reader may find the hypertext to be more of an advantage than the equivalent linear text.

Hypertexts also cause additional processing load by making the reader responsible for navigating the text. Skilled readers, who can process the text automatically will not have as much interference from the controlled processing task of having to make choices of where to go as they read. Readers with poor reading skills are using a lot more controlled processing for their reading processing and thus will likely have a greater amount of interference from the additional task of navigating the text. This will not let

them generate as many hypotheses about the text as they read it, making it harder to integrate the information.

Thus, the abilities of the reader may interact with the type of text format and how that text format is implemented. A well structured hypertext which does not impose a large navigational load on the reader may be an advantage for poor readers. On the other hand, poor readers may have great difficulty with a less well structured hypertext when compared to the equivalent linear text.

The Role of Reader's Strategies in Hypertext.

Examination of strategies in hypertext have thus far concentrated on the search strategies of readers when hypertexts have been treated as information retrieval systems. Since hypertexts provide additional navigational flexibility to readers, one could expect that readers would employ a variety of strategies. As evident from the Goldman and Saul (1990) study, readers use a variety of strategies in order to maintain coherence through a linear text. It is expected that readers of hypertexts would do the same, since they must still maintain coherence in order to develop a coherent mental representation of the text.

Part of the goal of developing hypertexts is to have mechanisms to simplify the navigational strategies of readers. Hypertexts were developed so that readers could get to the relevant sections in a more efficient manner than linear texts. For this reason, it is expected that the strategies of the reader can depend greatly on what mechanisms are implemented in the hypertext. A hypertext based primarily on search will tend to structure a reader's strategies to using search. However, these mechanism must also match the goals of the reader and these goals can influence the choice of strategies. Hypertext typically permits a reader to see more of the structure of the text and therefore it is somewhat easier to find the relevant places in the text. Thus, the reading strategies in hypertext could be expanded from those found in reading a linear text and could depend on the mechanisms provided in the hypertext and the goals of the reader.

Overview of this Research

Overall, there are a variety of factors that can cause differences in the readers' comprehension and strategies of a text written in hypertext and linear form. This research focused on just two of these factors, the coherence of the text and the goals of the reader. It was hypothesized that the coherence of the text will affect the comprehension, in that a hypertext which does not maintain as much coherence as the equivalent linear text will result in deficits in comprehension for readers who do not have background

knowledge of the topic of the text. A second type of hypertext was also tested which enforced additional coherence in the text in order to evaluate whether the lack of coherence in a hypertext could be improved. In addition, the goals of the readers were manipulated so that some of the readers had to find specific information while other readers read the text for general knowledge. It was hypothesized that manipulating the goals of the readers would change the reading strategies used. Readers with specific goals would find an advantage with the hypertext, since it permits fairly simple navigation to the relevant parts of the text. On the other hand, readers with general goals would find an advantage for the linear text, since it presents a single coherent set of information on the text.

CHAPTER V

EXPERIMENT 1

Rationale

This experiment investigated readers' comprehension and reading strategies of a linear text compared to two hypertext versions of a chapter from an undergraduate level economics textbook. The goals of the reader were manipulated so that half of the readers read the chapter for general knowledge, while the other half read the chapter in order to find certain specific pieces of information.

This design permits several comparisons. The first is whether there are differences in readers' comprehension and strategies between the hypertexts and the linear texts. It is hypothesized that since readers will have little background knowledge on the subject of the text, the lack of coherence in the hypertext may adversely affect the readers' comprehension when compared to the comprehension of the readers of the linear text. The second comparison is whether the goals of the reader interact with the format of the text. More specifically, do subjects use different reading strategies on the texts based on the goals, and does one format of the text afford better comprehension than another? It is hypothesized that hypertext may be more suited for tasks involving information search, so subjects with specific goals, in which they must find the relevant information, may perform better on the hypertext.

As described in the previous chapter, one of the hypothesized problems with hypertext may be that the jumps from node to node may not present as coherent a chapter as that of the linear text, resulting in lower comprehension. In order to investigate this, two hypertexts were developed. One hypertext used standard features found in hypertexts, while the second used those same features, but also automatically modified the text during reading in order to make it more coherent for the reader. In this second hypertext, as the subject jumped from node to node, the computer calculated what information the subject missed when passing over certain nodes. The missing information was then inserted into the text in order to make the jumps more coherent. The coherent hypertext allows us to compare two ways of implementing a hypertext in order to determine specifically whether adding coherence to the hypertext will improve the comprehension of the

text. This comparison is thus similar to that of Britton (e.g., Britton & Gulgoz, 1991), in that it has the goal of evaluating whether principled modifications to the text can bring about comprehension improvements.

One of the problems of comparing a hypertext and linear text using the same text is finding a text that can be used in both formats. A further constraint for the experiment is that the text must be a reasonable size for a short-term experiment (i.e., within one 2 hour experimental session), and at an appropriate level for undergraduate subjects. For some studies in hypertext, large texts (i.e. hundreds to tens of thousands of pages) have been used. Large hypertexts may be more appropriate for studying how people find information in the text. However, because the emphasis of this research is on the reading process rather than the search process, a smaller text was chosen.

At the time there are few existing hypertexts that fulfill the above requirements. Thus, a chapter from an undergraduate introductory textbook on economics was chosen as the experimental text. Although the text was originally in linear form, it fulfilled several requirements for both using as an experimental text and for conversion into a hypertext format. The text was well written, done by two of the more recognized economists in the field (Paul Samuelson and William Nordhaus). The book was in its 13th edition and thus, had a highly evolved structure and clearly defined sets of arguments. The organization of the text was in a hierarchical structure in which the authors also made several cross-hierarchical references. Therefore, the text was converted into a hypertext using guidelines from several works on hypertext (e.g., Martin, 1990; Nielsen, 1990; Shneiderman, 1987b). While it could be argued that this conversion of the text into the hypertext may not have been done in an optimal manner, the goal of this research was to investigate specific features of hypertext rather than the hypertext as a whole. For this reason, a minimum set of hypertext features were chosen in order to be able to abstract differences in the reading process without those features interacting with the many possible hypertext features.

The comprehension of the text was evaluated using several measures. When evaluating comprehension, there is a distinction between what people *remember* from the text, and what they *learn* from the text. What is remembered from the text indicates the quality of the reader's textbase, while what they learn from the text indicates the quality of the reader's situation model of the text. Typically, evaluations of both linear text and hypertext tend to examine one or the other of these features, although there may be large differences between the two. A reader may be able to recall information from the text, but not be able to use the information effectively. For this experiment, a variety of types of questions were used to evaluate comprehension. Some of the questions required subjects to recall information from the text, while other questions required subjects to use what they had learned from the text and apply it to new examples.

Method

Subjects

Eighty three undergraduates taking an introductory psychology class at the University of Colorado participated for course credit. The subjects were screened to ensure that they had not previously taken any courses in economics. Ten subjects were removed from the experimental analyses due to either failing to follow instructions or to equipment failure.

Materials

The third chapter from 13th edition of the introductory textbook *Economics* by Samuelson and Nordhaus (1989) was chosen as the text. The chapter, titled *Markets and Command in a Modern Economy*, was chosen because it was at an appropriate introductory level for undergraduates, was well written, and had a clear hierarchical structure. In addition, there were multiple places in the text where the authors referred to other sections in the chapter that spanned across the hierarchy. These places in the text were used to implement cross-hierarchical links in the hypertext format. The hierarchical structure of the chapter (without cross-hierarchical links) is shown in Figure 1.

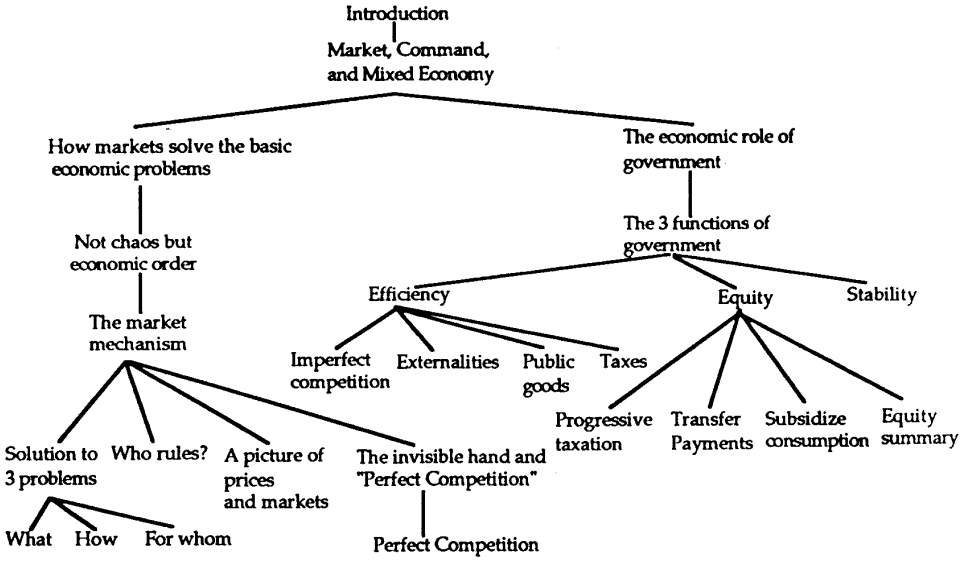


Figure 1. Hierarchical structure of the chapter

The text was edited to remove the final section of the chapter, all references to the linear order information (i.e. "the first section will discuss ... while the second section will discuss ..."), two figures, the references to those figures, and all references to future chapters. The resulting length of the chapter was 6018 words. The edited text had a Flesch reading grade level of 14.1. The text used is shown in Appendix A. The text was then converted into three text formats, one linear format and two hypertext formats, using Hypercard running on a Macintosh computer. For both the hypertext and linear text versions, subjects performed all actions by clicking on buttons on the screen with the mouse. The computer recorded the amount of time spent on each text section and all the actions made by the subject. The original font styles from the textbook chapter, such as bold and italic fonts were maintained for all versions.

Linear Text. The linear text version of the chapter was created to be similar to that of the original version of the text provided by the book, except that it was presented on the computer. The computer presented a picture of a page of a book and provides the subject four buttons to turn one page forward or back in the book, and to go to the start or end of the chapter. The initial page of the chapter contained the table of contents for the chapter in outline form. The total length of the chapter was 25 pages. One page of the linear text is shown in Figure 2.

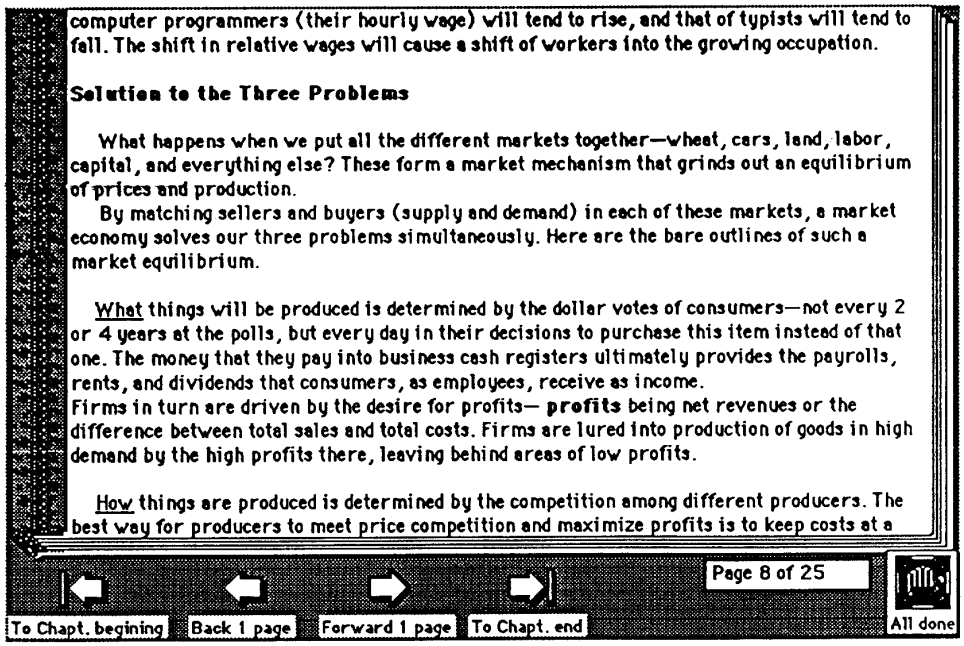


Figure 2. Page from the linear text version of the chapter

Hypertext. As discussed in Chapter II, there exist no definitive rules on how to develop a good hypertext. For this reason, it is difficult to develop an optimal hypertext for the economics text. Nevertheless, to develop a hypertext version of the chapter, principles that have been used in other hypertexts and have also been shown to be successful in hypertext and human-computer interaction research were used (e.g., Martin, 1990; Nielsen, 1990; Shneiderman, 1987b). The primary goal used in developing the hypertext was to maintain the structure and references used by the original authors. The structure of the hypertext was based on the hierarchical structure provided by the author's chapter outline. Each section of the authors' text was identified as an individual node. This resulted in a text consisting of 26 nodes.

One node of the hypertext is shown in Figure 3, illustrating the different features the subject could use. Readers of the hypertext were able to navigate along the paths of the hierarchy by either moving down the hierarchy by choosing nodes below the current one (buttons below the text), or moving up levels in the hierarchy ("Go up one level" button). The "Go to chapter beginning" button permitted the subjects to jump to the top level node of the text. Because the amount of text in some of the nodes was so large that the whole text could not fit into one screen, these nodes had scrolling buttons on the side of the text to permit subjects to scroll through the text. Clicking on the scrollbar buttons moved the text up or down one screen full.

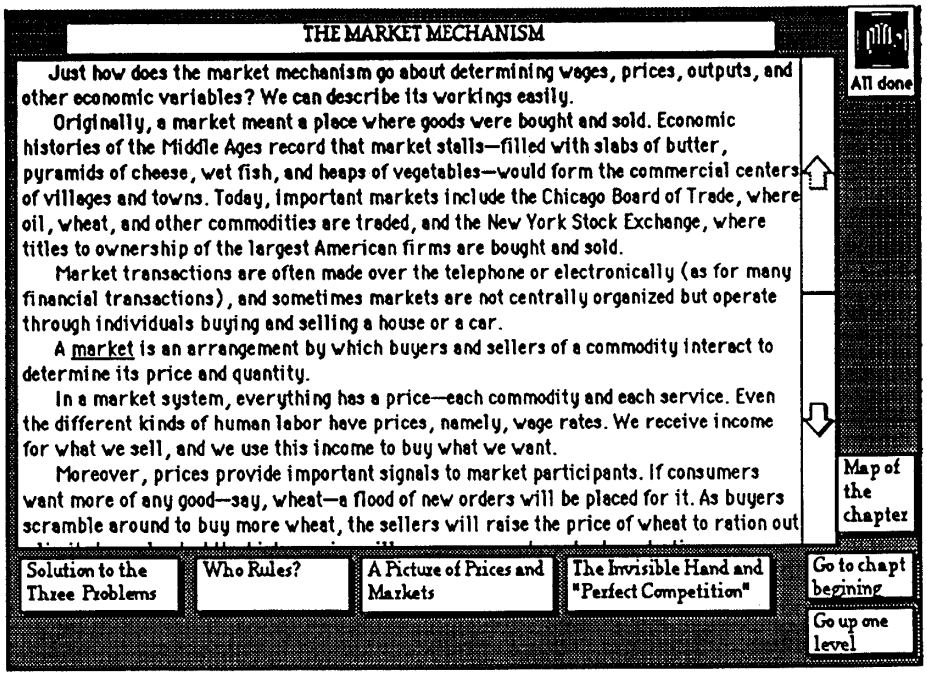


Figure 3. Node from the hypertext version of the chapter

In addition to the hierarchical connections, there were also connections between related text sections that the author had mentioned in the text. These cross-hierarchical connections were shown in the form of highlighted text. A subject could click on this text to be taken to the relevant node. There were 17 cross-hierarchical links in the text. The hypertext also provided a map of the hierarchical structure of the chapter in order to aid in navigation. This map was equivalent to the hierarchical structure shown in Figure 1. When looking at the map, subjects could click on any text node in order to go to that node. This permitted subjects to travel to any part of the text without using either the hierarchical or cross-hierarchical links.

Coherent Hypertext. A second version of the hypertext was created that was exactly the same as the first hypertext version, except that the first paragraph of each text node could change dynamically based on the previous actions of the reader. When using the map or the cross-hierarchical buttons in the hypertext, subjects could make jumps that would cross parts of the hierarchy into a very different context. However, this could result in difficulties for the subject, since this would not present a coherent flow of arguments in the text. For the coherent hypertext, the first paragraph was designed to provide necessary bridging information to make the current node cohere with the previous node read.

In order to determine what information should be inserted in these paragraphs, a macropropositional analysis of the text was first performed. For each text node, the primary macropropositions from the node were then put into sentence form. This resulted in texts of one to three sentences long that represented a summary of the primary information in each text node. When the subject read the hypertext, the sentences were inserted into the text, when necessary, using the rules shown in Table 1.

TABLE 1.

Rules for inserting coherence sentences

Type of jump	Coherence sentences to insert
Jump to child of current node	No sentences inserted
Jump to parent of current node or higher in hierarchy	No sentences inserted
Jump to sibling of current node	No sentences inserted
Jump to grandchild or lower of current node	Insert sentences from all child nodes that were jumped over.
Jump across any part of hierarchy	Determine single common ancestor of current node and new node and insert sentences from all nodes starting at the ancestor node down to the new node.

The sentences were added in any case where a subject would jump out of the current context in which he or she was reading. For this text, jumping out of a context was considered crossing part of the hierarchy or sub-hierarchy, or jumping down more than one level in the hierarchy. However, jumping to a child or sibling of a node would be considered within the same context, since readers would still have the appropriate background context to understand the new text. No changes to the text were made if the jumps were within the same context. The words "General Background" were inserted at the top of the inserted sentences in order to delineate these sentences from the rest of the text of the node. Because there are multiple ways to get to any text node, the general background paragraph of any text node changed based on where the subject was previously in the text. A sample node from the coherent hypertext with the additional background information paragraph inserted is shown in Figure 4.

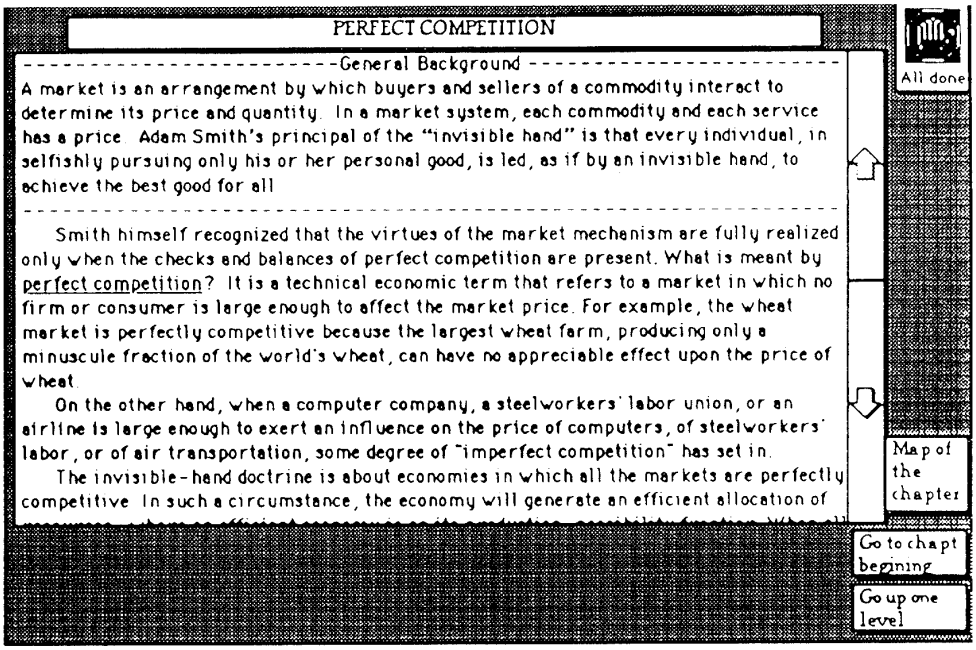


Figure 4. Node from coherent hypertext with context paragraph inserted.

Training materials. Most subjects were expected not to be familiar with the Macintosh interface and how to use the buttons necessary to navigate through the linear and hypertexts. An online tutorial was created that introduced each feature of the text format and then gave the subjects a sample text in which to practice using the text format. The sample text consisted of an encyclopedia article on Portugal that was 2145 words long using 14 sections in a hierarchical form. This text was converted into the three formats. For the linear format, this resulted in a text that was 11 pages long.

Questionnaire. After reading the text, the subjects' comprehension was measured using a variety of types of questions printed on three sheets of paper. These questions included one general essay question, five short answer essays, and eight multiple choice questions. The short answer essays and the multiple choice questions were chosen from multiple sections of the text in order to evaluate learning from different parts of the text. In addition, eight questions were given about the ease of using the text format and the amount of computer experience the subject had. The questions used are shown in Appendix B.

Design and Procedure

The experiment was a 3 by 2 between subjects experiment manipulating the format of the text and the reading goals of the subject. The three formats of the text were the linear version of the text, the hypertext version of the text, and the coherent hypertext version of the text. The reading goals of the subjects were manipulated by providing different instructions about what they should read for in the text. Half of the subjects were instructed that after reading the text, they would have to answer questions on a specific topic "the relationship between the invisible hand, perfect competition and imperfect competition". That topic covered several sections that occurred across different parts of the hierarchy of the chapter. The other half of the subjects were instructed that after reading the text they would have to answer general questions about the subject of the chapter "the role of the markets and government in the price functioning of an economy".

Subjects were randomly assigned to conditions. The order of the experiment was: (1) initial training in using the computer's text interface, (2) practice using the computer with the text interface on the sample text, (3) reading the text on the computer, and (4) answering questions about what they had read.

Before reading the economics text, all subjects were individually run through a tutorial on the Macintosh computer on how to use the features for the particular text format. In this tutorial, subjects were first shown each individual feature of the text format and were allowed to try it out. After seeing all the features, they were given a sample text on Portugal. They were instructed to practice using the features with the sample text although that they did not actually need to read the sample text. They were given three minutes to practice using the features before starting to read the actual experimental text. After reading the practice text, they were presented one screen of text and instructed to read the text and then press a button on the screen. The amount of time to read that screen of text was recorded in order to have a covariate reading speed.

Subjects were then given a piece of paper with either the specific or general instructions on what they should read for in the text. They were further told to read the text until they felt satisfied that they could answer questions on that topic, and when they were done, to click on the "All done" button on the screen.

After completing the reading of the text, all subjects were given the same set of questions. They were initially given a sheet of paper and instructed to write an essay on what the chapter was about. Once they had completed the essay, they were given the set of multiple choice and short answer essay questions. The questions included questions addressing the issues provided to the subjects in the specific goal condition, but also included questions from other sections of the text in order to evaluate learning from other sections.

After completing the comprehension questions, they then answered several questions about the ease of using the interface and the amount of experience they had using computers. Finally, before completing the experiment, subjects also signed a form to release their SAT scores so that they could be used as covariates for the analyses.

Results

The primary data collected were the time-stamped records of all the actions of the subjects as they went through the text. In addition, the answers to the multiple choice and essay questions were graded in order to obtain measures of the subjects' comprehension of the text and what parts of the text were understood. These data therefore permit analyses of the effect of the text format and goal manipulations on the subjects' navigation strategies and ease of comprehension of the text.

Analysis of Reading Times and Comprehension Measures

Reading Times. The total time to read the economics text was calculated from the time-stamped records. The mean time to read the text was 38.4 minutes with a standard deviation of 15.9. The reading times by condition are shown in Figure 5. There were no significant differences between the reading times based on the format of the text, $F(2,67)=0.27$, $p=.76$, or whether the subjects had specific or general reading goals $F(1,67)=2.55$, $p=.12$. However, since the lack of differences in reading speed may partially be due to the variability in subjects' reading speeds, a reading speed covariate was added. The reading time for the instructions for each of the three text formats was used. Since the instructions differed slightly based on the format given, the reading times were all converted to Z-scores. When this covariate was added, there was a significant difference in reading speed based on the subjects' reading goals, $F(1,66)=4.16$, $p<.05$, with the subjects who were given the specific reading goals taking a shorter amount of time. This indicates that subjects spent less time reading when they had to find specific information than when they had to read the entire text.

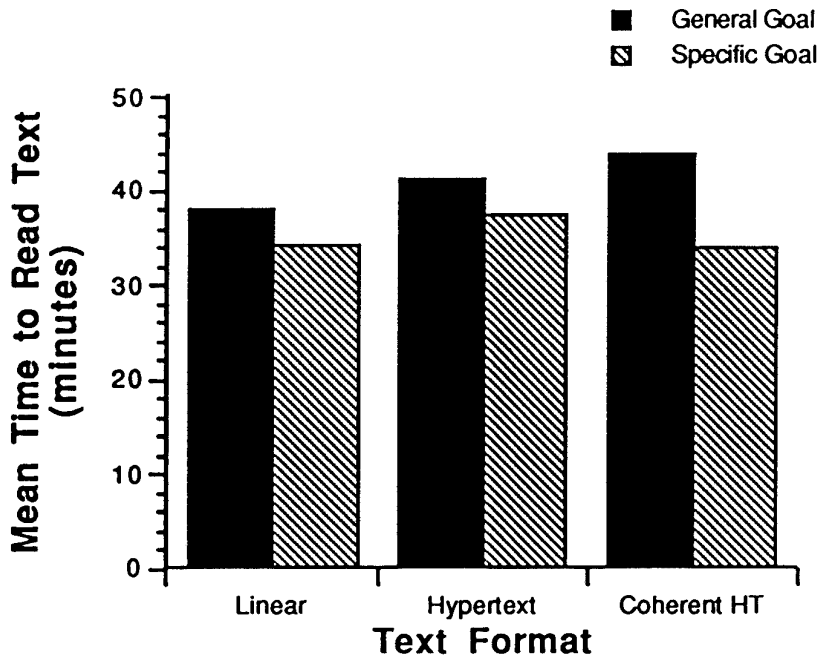


Figure 5. Mean reading time for the three text styles and two reading goals.

Comprehension Measures. Two types of measures of comprehension were obtained. The first measures of comprehension was calculated through grading the multiple choice and short answer essay questions. The second type of measure of comprehension was obtained through performing a propositional analysis of the subjects' essays.

The multiple choice and short answer essays were graded and four scores were calculated for each subject. These scores were created by grouping similar sets of questions in order to rate comprehension of specific features of the text or text interface. The scores were: total score on all questions (13 questions), score on questions about the content of the text (10 questions), score on questions about the structure of the text (3 questions), and score on questions about specific goal topic (5 questions). The mean percent correct for the four comprehension scores over the different text formats and reading goals are shown in Table 2.

TABLE 2

Mean percent correct for the four comprehension scores

	Specific Reading Goals			General Reading Goals		
Comprehension Measure	Linear Text	Hypertext	Coherent Hypertext	Linear Text	Hypertext	Coherent Hypertext
Overall Score	66.7	63.1	63.8	56.4	66.2	66.8
Content Score	71.3	71.1	69.2	58.8	71.4	68.8
Structure Score	36.0	48.3	34.7	32.0	58.3	59.0
Specific Score	74.2	81.0	73.4	55.0	63.6	55.0

Overall, the mean scores were not that high, indicating that subjects did not perform very well on these comprehension measures. However, all scores were well above the chance level. There were no significant group differences for the total score and the content score. The addition of covariates (SAT scores, amount of computer experience, and reading speed) did not change these analyses. Thus, the test questions showed no significant differences in the overall comprehension of the text.

While there were no differences in comprehension, there were differences in the subjects' knowledge of the structure of the text. The mean structure scores for the three text styles and two reading goals is shown in Figure 6. Subjects with the general reading goal had better understanding of the structure of the text than those with the specific reading goal, $F(1,67)=4.01$, $p<.05$. Subjects reading the hypertexts also had a better understanding of the structure of the text than those reading the linear text, $F(2,67)=5.07$, $p<.01$. This effect may be due to the fact that subjects in the hypertexts could use the map that was provided with the hypertext, while subjects in the linear text relied on a table of contents for navigation. The text style by goal interaction was marginally significant, $F(2,67)=2.73$, $p<.08$. Figure 6 shows that for the subjects with the linear text format, there is essentially no difference in the knowledge of the text structure based on the type of goal. This likely indicates a floor effect. While subjects in the two hypertext text formats perform better when given the general reading goals, subjects in the coherent hypertext with

specific goals performed only about as well as the subjects with the linear text. This likely indicates that subjects in the coherent hypertext could rely on the coherence of the text to help them navigate to the relevant places in the text without as much use of the map.

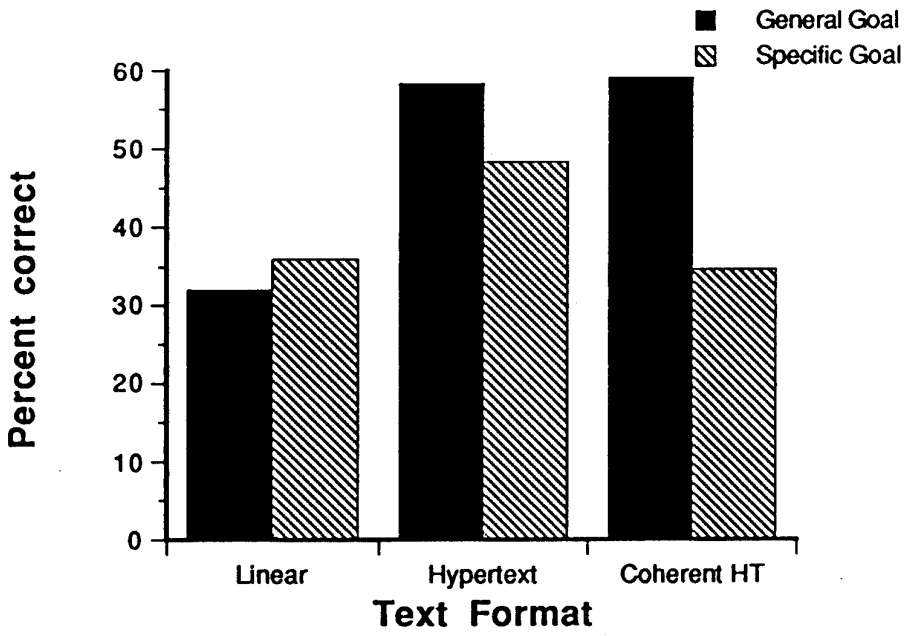


Figure 6. Mean percent correct questions on the text structure for the three text formats and two reading goals.

An analysis of the scores on the questions covering the specific goal topic found that subjects who were instructed to read for information on that topic scored higher than those who read the text for general knowledge, $F(1, 67)=8.46$, $p<.01$. This indicates that manipulating the goals of the subjects did result in somewhat different reading comprehension and those who were instructed to find the specific information were better at retaining that information.

While the multiple choice and short answer essays provide measures of certain pre-chosen questions, full essays provide a better description of the information retained and deemed most important by the subjects. The essays written by the subjects were analyzed by initially identifying all the primary macropropositions from the economics text. Eighty-nine macropropositions were identified from the text. Each essay was then coded by identifying which of the macropropositions were used in the essay. The mean number of macropropositions in the essays was 9.1 with a standard error of the mean of

0.55. There was quite a bit of variability in the quality of the essays with the number of unique macropropositions ranging from 1 to 26.

One indication of the quality of the essays would be the number of these relevant macropropositions generated. It would be expected that better essays would have a greater number of these relevant macropropositions. There were no significant differences in the number of macropropositions generated based on the different text formats, $F(2,67)=.06$, $p>.93$. There were also no differences for the number of macropropositions generated based on the subjects' reading goals, $F(1,67)=1.3$, $p>.25$.

An analysis was also performed on the hierarchical level of the propositions that were generated. This permitted a comparison of whether certain of text formats or goals afforded better overall understanding of the higher-level concepts or of the lower-level concepts from the text. The macropropositions were coded based on the level they occurred in the text. Based on the hierarchical structure of the text, four levels were identified. The top-most level represented propositions taken from the first two nodes of the text, while the bottom-most level represented propositions taken from any of the nodes on the bottom leaves of the hierarchy. The other two levels were represented by macropropositions from the upper and lower portions of the intermediate nodes.

For each subject, the percent of macropropositions from each level was calculated. Overall, at each of the four levels, there were no significant differences in the percent of macropropositions due to the text format. There were also no differences in the percentage of top-level macropropositions generated based on the goals of the reader. However, readers with the specific reading goals generated a higher percentage of macropropositions from the bottom-level nodes, $F(1,67)=4.99$, $p<.05$, while readers with general goals generated a higher percentage of macropropositions from the second and third levels, $F(1,67)=13.44$, $p<.01$ and $F(1.67)=7.5$, $p<.01$. Since most of the specific information that subjects were to find in the specific reading goals condition were in the bottom level nodes, this results indicates that those subjects were better able to recall that information. Overall, the propositionalization of essays showed that while the subjects' comprehension did differ somewhat based on the instructions to the subjects, comprehension did not differ based on the type of text format.

Analysis of Readers' Strategies and Coherence

The analysis of comprehension and reading times above shows that there were very few differences between the text formats. One possible explanation of the lack of differences may be found in examining the strategies used for reading through the text. If subjects used similar strategies in reading all three text formats, then they would all see approximately the

same text and in the same order. One would then not expect differences in reading comprehension between the text formats. In addition, an analysis of readers' strategies would permit a determination of whether the goal manipulation and the text formats caused different reading strategies to be used by the subjects.

Analysis of the Amount of Text Read. One analysis of strategies is to examine statistics on the amount of the text that the subjects actually read, and what methods they used for getting to the different parts of the text. Figure 7 shows the mean percentage of the nodes or pages that the subject actually read. For this analysis, a page or node was considered "read" if a subject spent more than 5 seconds on that page or node. Figure 7 shows that with general reading goals, subjects in all three text formats tend to read almost the entire text compared to subjects in the specific reading goals, $F(1,67)=18.5$, $p<.01$. This is consistent with the reading time results presented earlier. Subjects in the specific goal condition took less time to read the text because they read less of the text. There was also a significant difference between the text formats $F(2,67)=6.0$, $p<.01$. Subjects given the linear text read almost the entire text (93%), while those in the hypertext read fewer text nodes (86% for the hypertext, 72% for the coherent hypertext).

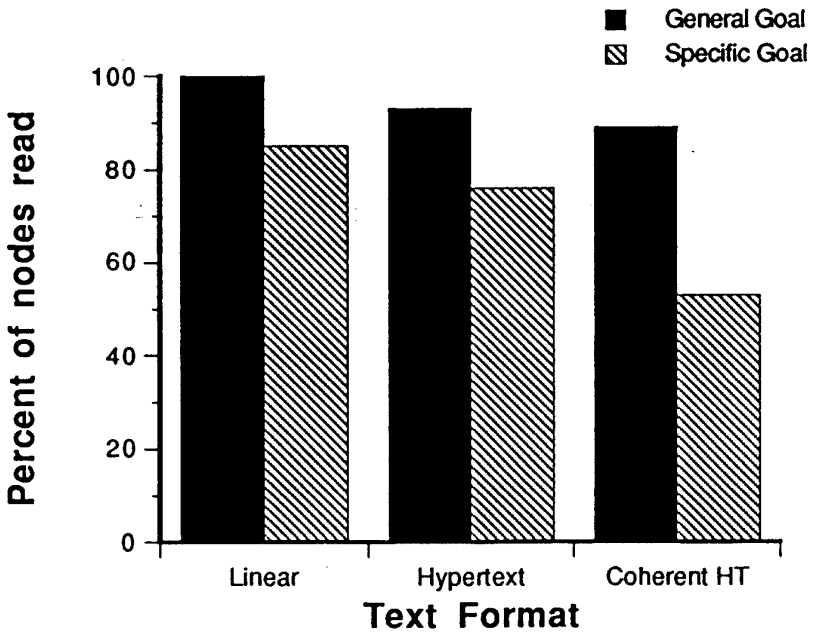


Figure 7. Mean percent of nodes or pages read for the three text formats and two reading goals

The fact that the subjects who were given the linear text read almost the entire text even when given specific goals for finding information may partially be an artifact of the text format. Since subjects were restricted to only moving forward and backward one page in the linear text format, they would tend to hit every page of the text. It is interesting to note that subjects in the coherent hypertext condition read the least number of nodes. Since the coherent hypertext provided additional background information, subjects may not have felt that they needed to read to each node and could rely on the background context paragraphs to help guide them.

Order of the Information Read and the Role of Coherence. One of the hypotheses of this research was that the order that the information was read would affect comprehension. Subjects who read the text in a non-coherent manner would not have as great comprehension of the text as those who read it in a linear manner. Two statistics were computed to represent the coherence of the transitions made by each subject when going from one node or page to another. For both of these statistics, the transitions were only counted if a subject spent more than 5 seconds on a particular page or node. This permitted calculations based on only the nodes or pages that the subjects had actually read. Thus, transitions by the readers of the linear text could be counted as non-linear if they skipped over pages by spending less than 5 seconds on any particular page.

The first statistic calculated was the percent of linear transitions made by the subject. A linear transition was a transition from one node or page to another that was chosen in the same order as the original text. Thus, reading the linear text in the linear order or reading the hypertext in a depth-first manner going down through the hierarchy starting from the left would result in 100 percent linear transitions.

The second statistic calculated was the percent of coherent transitions made by the subject. A coherent transition was considered a transition from one node or page to another in which both nodes or pages were still within the same context. This would included such cases as jumping to a parent, child or sibling of the current node. Thus, all children of a node would be considered to be within the same context and they could be read in any particular order. The rules for inserting the coherence sentences from Table 1 were used to determine whether the transition was coherent. Based on these rules, for any transition where no coherence sentences would be needed to be inserted, then the transition was considered coherent.

An analysis of the relationships between the coherence measures and the comprehension measures showed few significant correlations. There was a marginally significant correlation between the percent of coherent transitions and the number of propositions generated by subjects in the essays ($r=.21$, $p<.07$). This suggests that the more coherently the text was read, the more relevant propositions were generated for the essay. These results

therefore provide some evidence that the coherence of the text was related to the comprehension of the subjects.

While evidence for relationships between the coherence measures and the comprehension were not very strong, it is important to look at the effect of the text format and goal manipulations on the two coherence measures. Figure 8 shows the mean percent of linear transitions for the three text formats and two reading goals. There was a significant difference between the number of linear transitions in the linear text and the two hypertexts, $F(2,67)=35.8$, $p<.01$, however, subjects reading the hypertext still made approximately 55 percent of their transitions in the same linear order as the original text. This suggests that subjects in the hypertexts were primarily using a depth-first strategy for moving through the hierarchy. The interaction between the text format and the reading goals was not significant, indicating that the reading strategies did not vary with the type of text format.

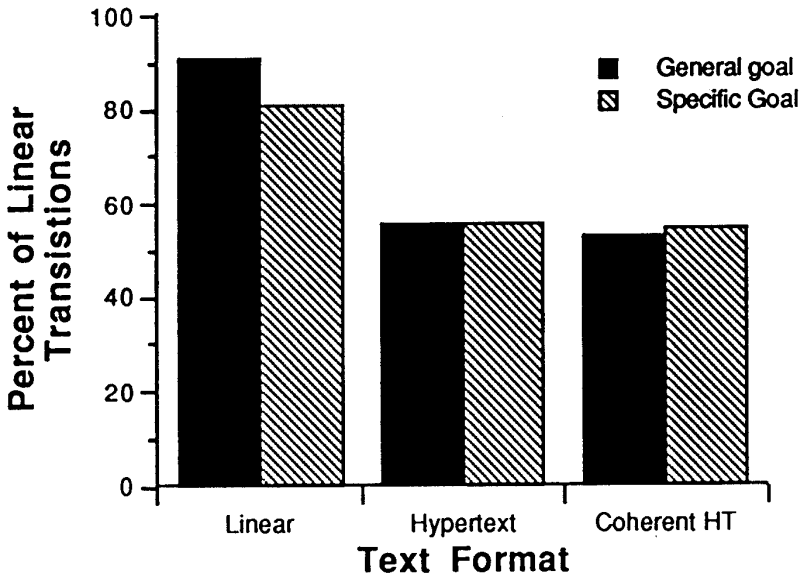


Figure 8. Mean percent of linear transitions for the three text formats and two reading goals

Figure 9 shows the mean percent of coherent transitions for the three text formats and two reading goals. For this coherence measure, there were no significant differences between the text formats or the reading goals. For all groups, subjects made 80 to 90 percent of their transitions in a coherent manner. Although this is not surprising for readers of the linear text, it

suggests that readers of the hypertexts only made transitions within the same context and seldom used the cross-hierarchical buttons or map buttons to move them to very different areas of the hierarchy.

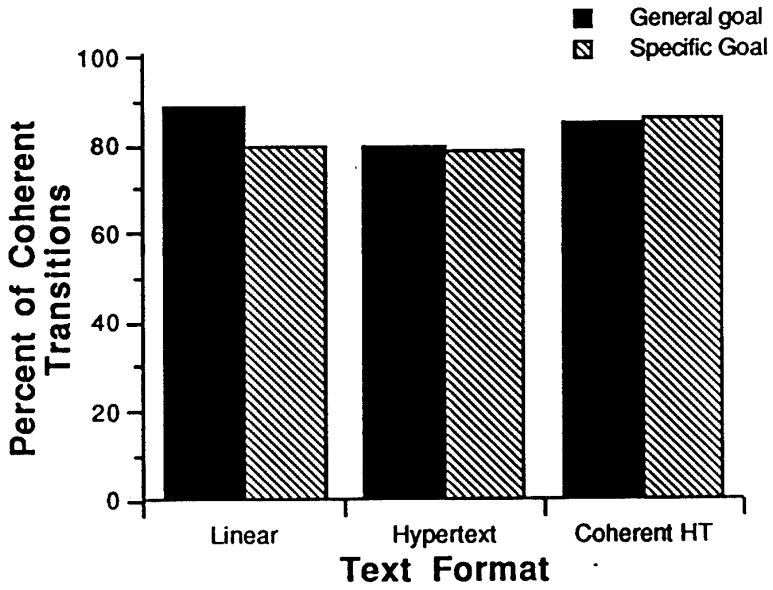


Figure 9. Mean percent of coherent transitions for the three text formats and two reading goals

The fact that there were no group differences in the coherent transition measure suggests that the lack of differences between the three text formats from the comprehension measures could be due to ceiling effects. All subjects were seeing approximately the same text in about an equivalently coherent manner, and thus their comprehension would not differ greatly. Similarly, the lack of correlations between the subjects' strategies and the comprehension measures could be due to the fact that there was just not enough variability in the strategies used by subjects, since they tended to use a very rational approach of maintaining coherence as they navigated through the text.

Hypertext Strategies

The above results show that the hypertext readers used features in the hypertext that permitted them to move through the text coherently. This raises questions of what types of strategies did the hypertext readers use. An

examination of the time-stamped record of their actions permits an analysis of what navigational features the hypertext readers relied on to move through the text.

The percent of usage of each of the five hypertext features for each subjects was calculated. Table 3 shows the breakdown of the percentage of navigation methods used by readers of the hypertext and coherent hypertexts combined The "hierarchical down" buttons and the "go up one level" buttons allowed only hierarchical movements through the text, which therefore permitted only coherent transitions. Overall, these two buttons made up 64% of total movements that the subjects made. Nevertheless, considering that overall subjects made 80 to 90 percent of all transitions in a coherent manner, subjects must have used other navigational features in addition to the hierarchical buttons to make additional coherent transitions.

TABLE 3

Mean percent of usage of the hypertext navigation features

Navigation Method	Percent of usage
Map	26.4
Hierarchical down buttons	39.6
Go up one level	24.6
Hot word (cross-hierarchical)	6.4
Go to top level	3.3

Navigating using the map and the "hot words", which gave hierarchical cross links, permitted non-coherent transitions. Thirty three percent of the transitions were performed using these methods. Subjects relied on the map for making about 25% of their transitions. Using the map permitted subjects to go anywhere in the text, and therefore, they could choose whether or not to make coherent transitions. Of all the transitions made by subjects with the map, 60% of these were coherent. Thus, although the map gave more flexibility of choices of where to go, subjects still relied on making primarily coherent transitions when faced with those choices.

As additional evidence that the subjects used a rational approach to navigating the hypertext versions in a coherent manner, we can examine the total possible ways of navigating the chapter. Given the hierarchy of this chapter, there are 995,328 possible orders in which to navigate the hierarchy in a coherent manner going to each of the 26 nodes once. This is based on the possible orders of following the hierarchy, ignoring the cross-hierarchical

links. However, with 26 nodes, there are a total of 26!, or 4.03×10^{26} possible orders to read those nodes. From this, the proportion of coherent orders through the hypertext would be only 2.47×10^{-21}. Thus, the set of coherent paths through the hypertext represents a very small subset of the total number of possible paths through the hypertext. The fact that subjects primarily relied on these coherent paths indicates that they were not just making random choices of the order, but that they used some way of guiding themselves through the text in a coherent manner.

Repairs to non-coherent transitions. Although the overall number of non-coherent transitions made by the hypertext were fairly small, it is important to know whether the transitions resulted in getting the subjects to the correct place in the text. Since verbal protocols were not used in this study, it is difficult to know whether the subjects were satisfied when they jumped out of one context and into another, or whether it resulted in some confusion. The time stamped record however, provides a measure of whether subjects remained in that new context or immediately jumped back into the previous context. Thus, when readers of the hypertext deviated from the coherent path, did they immediately tend to return to the context of the original path?

An analysis of the number of times that subjects made a non-coherent transition by jumping into a new context and then, on the next transition, jumped back to the original context was performed. Overall 27 percent of all non-coherent transitions resulted in the subjects immediately repairing the jump by returning to a node that was in the original context. Thus, even though there were few non-coherent transitions made overall, when subjects made non-coherent transitions, one quarter of them were immediately repaired. This again appears to indicate that subjects preferred to read the text within one context before reading parts of the text that occurred in a different context.

However, an alternative explanation of the repairs made by the subjects could be from cases when subjects chose to make a non-coherent transition to look up one specific piece of information in the other context and then once they had found that information, they jumped back to the original context. This would be akin to using the hypertext jumps to reference a piece of information to help clarify information in the current context. Without access to the subjects' actual motivations, it is difficult to clarify whether the repairs were made after the subjects accidentally jumped out of the context, or whether it was done intentionally.

Discussion

Timing and Comprehension Differences

Based on the comprehension measures, manipulating the reading goals of the subjects affected some of what the subjects retained from the text. Subjects with specific reading goals read less of the text and took a shorter amount of time to do the reading than those with the general reading goals. These subjects were also better at recalling more of the specific and low level information in the text, but were not as aware of the structure of the text. These results are not surprising. Telling people to find different information in the text, caused them to look for the information and therefore produced slightly different comprehension and strategies. However, the reading goal manipulation did not interact with the type of text format used. Thus, subjects were faster in all text formats when they had to find the specific information, and their reading strategies were not different across the three text formats.

Differences were found in the comprehension and reading time due to the reading goals, however no differences were found in the amount of time to read the text and few differences in the various comprehension measures used due to the text format. Although there were individual differences in the subjects' reading speed and the subjects' cognitive abilities that could affect these measures, using reading speed and test score covariates did not significantly change these results. The fact that there were few differences in the amount of time subjects took to read the text may also be a function of implicit demands of the experiment. Subjects were aware that they would receive two credit hours for the experiment, and they may have structured their time in reading the text with this in mind. If subjects had been given more freedom in reading, such as using the text for studying for an actual test or for reading for a class, the subjects may have developed somewhat different reading times and strategies based on the format of the text.

The comprehension results indicate that the three text format produced equivalent amounts of comprehension for the subjects. Subjects reading the hypertexts did display a better knowledge of the structure of the text and this is likely due to the fact that they had a map which presented a better visual representation of the structure than the table of contents that was used with the linear text. The coherent hypertext similarly did not improve the comprehension of the subjects over that of the hypertext or the linear text. It may be that the coherence sentences inserted were not the appropriate sentences that the subjects needed in order to avoid making bridging inferences. It may also be possible that, since the text for the coherence sentences was somewhat segregated from the main text of each node, it was ignored by the subjects.

It was expected that the hypertext would cause the readers to read the text in a less coherent manner and this would adversely affect the readers' comprehension. In addition, it was expected that the coherent hypertext version would improve the comprehension through providing the necessary bridging inferences that would have to be made when a reader made a non-coherent jump. These expectations were not confirmed by the results.

Readers' Strategies

The fact that the subjects' general comprehension of the text was equivalent was counter to the original hypotheses of this study. These lack of differences can be accounted for by looking at subjects reading strategies. For all three of the text formats, subjects tended to read the text in a coherent manner. For the linear text, almost all the transitions made by the subjects were in the linear order. Thus, even with the specific reading goals, subjects tended to read the text linearly, rather than searching through the pages to find the relevant information. Even though the hypertexts were not organized with an explicit linear structure, subjects still made approximately 55% of their transitions in the same order as the linear text. This indicates that subjects must have used heuristics to determine a coherent order through the text.

Although the linear coherence provides a measure of how similar the order the text was read to the original linear text, the topical coherence may provide a better measure of the coherence of the text. For this text, the children of a node are all approximately as relevant and follow on as coherently once the upper-level node has been read. Thus, jumping to any of the children will present a coherent transition. In all conditions, 80 to 90 percent of the transitions were made in a coherent manner which indicates that even when the a text can be easily read in a non-coherent manner, subjects tended to choose the coherent paths through the text. While subjects used the hierarchical features provided with the hypertext, they also used the map feature for navigating over 25 percent of the time. Although the map permitted subjects to make non-coherent transitions, when using the map, they nevertheless primarily chose to make coherent transitions. This indicates that subjects did not just rely on the hierarchical button of the interface to guide them through the text, but made reasoned choices of where to go. In addition, even when subjects made non-coherent transitions, over one quarter of them were immediatly repaired by returning to the original context.

Subjects were therefore displaying the rational behavior of determining the best method for reading through and comprehending the text. Because the subjects all saw the text in an equivalent coherent manner, their comprehension of the text would also be approximately equivalent. Therefore, it is not surprising that there were few differences in the

comprehension measures. These strategies may also explain why the coherent hypertext did not bring about any differences in comprehension from the original hypertext. In the coherent hypertext, the necessary bridging inferences were only inserted as "general background" when the subject made a non-coherent transition. Since almost all the transitions were coherent, this background information was seldom presented to the readers.

The strategies used by the subjects appear to be similar to those identified by Goldman and Saul (1990) in their strategy competition model. Their model posits a variety of strategies for maintaining local and global discourse coherence as subjects navigate through a text. Although their analyses looked at coherence at the sentence level, subjects in this experiment created coherent paths at the node level, where a node consisted of one or more paragraphs. The results of this experiment show a set of similar strategies. Subjects proceeded through the hypertext, and at decision points, chose to move to nodes that would help maintain the global coherence of the text.

While subjects used a rational approach for getting through the text in a coherent manner, it is not clear what constraints of the text were used to help guide them through the text. A variety of signals contained in the text could play a role in guiding subjects. For the linear text, it was likely that the subjects merely followed the linear order of the text and produced the most coherent path. This would be consistent with their past experiences with text.

However, since subjects had no previous experience with hypertexts, they would not be as familiar with what signals were available. Thus, it is not clear what information they could have used as signals of coherence. One of the primary signals could be the hierarchy of the chapter that was presented as a map and implicit in the hierarchical navigation buttons. This would impose some constraints on moving up and down the hierarchy coherently. However, the hypertext did present many other possibilities to make non-coherent transitions through using the map buttons and the hierarchical cross-links. Nevertheless, subjects appeared to avoid using these features in favor of the features that would get them through the text coherently.

Thus, this experiment shows that the strategies used by subjects in the hypertext were similar to those found in readers an equivalent linear text. Nevertheless, further research should examine what text signals the subjects used for navigating the hypertext coherently and what the reasoning the subjects had behind choosing their strategies. This information would help characterize what features of the hypertext simplify the reading process. The second experiment investigated these questions.

CHAPTER VI

EXPERIMENT 2

Rationale

The results from Experiment 1 showed that, although there were little differences in comprehension due to the three text formats, the reading strategies used by the subjects were all very similar. This indicates that subjects used heuristics to employ information from the text to guide them through the text in a coherent manner. While Experiment 1 showed evidence of these strategies, the experiment did not provide data that showed the reasons why the subjects used these strategies and what type of information from the text helped them to use these strategies. Therefore, the goals of the second experiment were to determine what lead subjects to maintain coherent paths through the hypertext and to determine what signals in the text helped the subject find these coherent paths. Because there is some amount of cognitive effort involved in using strategies for choosing the coherent path through the text, this second experiment could be conceptualized as an investigation into the problem solving that the subjects used when navigating the hypertext.

In order to have access to the subjects' strategies, verbal protocols were collected as the subjects used the hypertext (e.g., Ericsson & Simon, 1984). When reading a hypertext, much of the reading process will be automatic and therefore not accessible to the verbal reports. However, when subjects reach points in the text when they must make decisions about where to go in the text, the subjects would use controlled processing. These episodes should therefore be accessible to the subjects and could be expressed through their verbal reports. Thus, this experiment can help corroborate evidence from Experiment 1 and also provide qualitative information about the subjects' strategies and what information from the hypertext helped the subjects navigate in a coherent manner.

Method

Subjects

Six undergraduates taking an introductory psychology class at the University of Colorado participated for course credit. The subjects were screened to ensure that they had not previously taken any courses in economics. One subject was removed from the analysis due to an equipment failure.

Apparatus

The same apparatus as that used in Experiment 1 was used for this experiment. In addition, a microphone and tape-recorder were placed on the table next to the computer in order to record subjects' verbal protocols.

Design

This experiment used two of the conditions from Experiment 1. In both conditions, subjects read the hypertext economics chapter. Two of the subjects were given instructions to read the text so that they could answer questions on "the relationship between the invisible hand, perfect competition and imperfect competition" (specific reading goals condition). The remaining four subjects were given instructions to read the text in order to answer general questions on the topic of the chapter, "the role of the markets and government in the price functioning of an economy" (general reading goals condition).

Procedure

The procedure was the same as Experiment 1, except that subjects were asked to speak aloud as they read through the text. All subjects were run individually. Subjects were first run on the tutorial that provided training on using the features of the hypertext. They then received practice using the hypertext features using the short hypertext on Portugal. Once they had completed the hypertext training, they were given practice in providing verbal protocols. For this, the subjects were given tasks such as counting the number of windows in their home while thinking out loud as they performed the task. After they had completed two verbal protocol practice tasks, they were then told to read the hypertext and were given either the specific or general reading instructions. The subjects were further instructed that during the experiment, they were to say what they were thinking at all points as they read through the text, although they were to read the text

silently. The experimenter stayed in the room with them throughout the experiment. If a subject was quiet for any long period of time and did not appear to be reading, the experimenter prompted the subjects to say what he or she was thinking at that point. Subjects were instructed to read the text until they felt they were ready to answer questions on the topic. Although subjects were told that they would have to answer questions about the text after reading the hypertext, subjects were not given any questions about what they had read, but instead were asked to provide retrospective reports on any strategies they may have used for reading the text.

Results

Two sets of data were used in the analyses. The first set of data was the time stamped records of all the actions made by the subjects as they read through the hypertext. This permitted an examination of the order that the text was read. The second set of data was the verbal protocols provided by the subjects.

Overall, the verbal protocols were fairly sparse. Subjects spoke very little during the experiment and tended to express what they were doing rather than what their reasoning was behind what they were doing. This likely indicates the difficulty of combining the silent reading process while verbalizing aloud about what one is thinking. Because of the sparse verbalizations, it was not possible to make a full categorization of the different behaviors expressed in the verbalizations. Nevertheless, the verbal protocols do provide some qualitative information about reading strategies used by the subjects.

Overview of Reading Strategies

Based on the time-stamped record of the actions taken by the subjects reading the hypertext, the paths chosen by the subjects were very similar to those used by subjects in Experiment 1. The subjects tended to read the text in a very coherent manner, seldom jumping into different contexts. Of the subjects who were given the general reading goals, two primarily used a depth-first search method through the whole hierarchy, going to all the nodes. The third subject first used one cross-hierarchical link, then went back to the top level and did a depth-first search through only the upper levels of the hierarchy, reading 65% of the nodes.

The subjects who were given the specific reading goals read only about 40% of the nodes. Thus, they were much more selective about what they read in the text. For both subjects, they initially used the map to identify the nodes in the text that would contain the specific information that they would need to read. However, once they had located this information in the map, they

then read down through the hierarchy to that information. Thus, although they had found the relevant information, they still used coherent paths to get down to that information.

Subjects' Verbalizations About Reading Strategies

From the time-stamped record, we see that subjects still followed coherent paths through the text, even when they needed to get to a specific node. The verbal protocols provide some additional evidence about why they used these coherent paths. Two of the subjects used the cross-hierarchical links near the beginning of their reading and ran into confusion about where they were in the text.

> Umm, I'm thinking, I want to go up one level, and see what was up there and if I'm still lost I'll go to the map and check it. completely lost ... just, I want to go back , I'd like to go back to where I kinda started with. I think there are three ways the market, gets involved, equity, stability, or something like that, and I'm just in another chapter now. (subject hg3001)

Jumping across the hierarchy moved them into a different context. Although the current node that they were reading was slightly related to what they were reading previously, most of the text nodes around that node were on a fairly different topic. Because of this, the subject is trying to get back to the previous context about the "ways the market gets involved".

The fact that the cross-hierarchical links were not used that much likely indicates that subjects were aware that they primarily wanted to stay within the same context. Subjects also expressed the fact that changing levels would change the context and instead opted for using depth-first search for moving through the text.

> I'm just going to go back and see if this is the only one. Rather than keep on going on this, or go in a level order. I think it is probably smarter to go on down the line. (subject hg3006)

When asked at the end of the experiment if he had used any reading strategies, the subject replied:

> Um, not really, I just went down, all of the (garbled) and whenever I got to the bottom, I just went to the next one over.... I felt like if I, say if I start at the top and it branches off to a subject and that subject branches ... you start at the top and that branches off to two subjects and each of those branches down. If I went to one, and then I went to the other, these two down here, below wouldn't make as much sense. I kinda felt that if you went, read to more, makes more sense. (subject hg3006)

Thus, the subject argues that using the depth-first method of proceeding through the text would allow the text to make the most "sense". In addition to an awareness of following the hierarchy, one subject also expressed the

strategy of covering all the text in one part of the hierarchy before moving on to the other parts of the hierarchy.

> I think I'm done in government. I've covered all, covered that. Uhuh, I'm going to go to... how markets solve the basic economic problems. (Subject hg3001)

This strategy has somewhat of the same effect as the depth-first strategy. By reading one part of the hierarchy before moving on to other parts, the subject avoids moving out of the current context at any point. Indeed, if this strategy is applied recursively, the two strategies are equivalent.

Subjects with the general reading goals used reading strategies for covering the whole text in a coherent manner. On the other hand, the subjects with specific reading goals tried to maintain the context around the specific information they needed to find.

> Um, I was mostly looking at the general topics and trying not to skip, go all the way down to the, that in case there might be other topics to the left that were on the same kind of hierarchy, or whatever. I tried to look at, get an overview of those other three before I went over to there, in case they had something to do with that [....] I just think I needed an overview of the little things around it in order to get at the answers to that question. (subject hs3004)

Thus, it was not sufficient to find the relevant information in the text. In order for that information to make sense, the subject needed background context that could be provided by reading the text from the nodes that were around the relevant node. In fact, based on the time-stamps, the subjects read most of the parent nodes of the relevant information before reading relevant information. Overall, the subjects' protocols indicated that they did employ strategies for staying within the same context. When they found themselves out of what they thought was the current context, they also expressed confusion, indicating their need for maintaining the appropriate context.

What Guided the Subjects' Strategies?

The verbal protocols show that subjects used strategies for maintaining the coherence of the text. It is also important to know what information they used to help guide them to use these reading strategies. From both the verbal protocols and the time-stamped record of actions, subjects seemed to rely primarily on the map to guide their strategies. Subjects consulted the map an average of 8 times throughout the session. This indicates that they used the map often to help orient themselves. The subjects' protocols also showed that they relied on the map for guidance.

> I'm thinking umm, I'm trying to decide whether to go back to the map of the chapter and work my over actually to keep working my way down the map. (Subject hs3004)

In addition to the verbalizations about the map, subjects also appeared to use the titles of nodes in order to guide them. One subject identified the three nodes that were under the node titled, "The three functions of government", and decided he had better cover all three functions. Thus, while the map of the hierarchical structure appeared to be the primary piece of information used to guide the subjects, the titles could identify relationships between items in the hierarchy and allow the subject to navigate based on following the labels.

Discussion

This experiment corroborates the results found in Experiment 1. Subjects used heuristics for maintaining the coherence of the text as they navigated through the hypertext. The verbal protocols showed that the subjects were aware of their strategies to read the information around one area of the hierarchy before moving on to other areas. In this manner, their strategies allowed them to always stay within the same textual context. While they expressed these strategies, their protocols did not show evidence of awareness of why they used those strategies as opposed to others. Instead, this depth-first approach to reading just "made sense".

When subjects jumped out of the current context, they expressed confusion. This result suggests that this can be a problem, in that subjects must make bridging inferences between the original context and new text in order to understand the new text. If this is not possible, they may face losing the original context. In addition, if they wish to restore the original context, they must use problem solving to find the correct action to return them to the part of the text with the original context. With hypertexts, this loss of context has often been described and is typically called being "lost in hyperspace".

The strategies used by the subjects show evidence that reading the hypertext is not just a reading process, but also a process of problem solving. Subjects developed heuristics for dealing with an unfamiliar text domain and text format. These heuristics also seemed to be based on the goals of the reader. For subjects with general reading goals, they read through the hierarchy in the most coherent manner possible for getting through the complete text. For subjects who were reading with the goal of finding specific information, they first located the information on the map. Once they had identified the relevant nodes, they then used the hierarchy to navigate down to the specific information, reading all the upper level nodes and then reading all information in proximity to the specific information in order to get a better context. Thus, the subjects developed problem solving strategies for developing the best representation of the text for their particular goal.

Some of the implications of these problem solving strategies to hypertext navigation will be described later in the general discussion.

Although subjects employed strategies for getting them through the text, the verbal protocols provide some evidence of what the subjects used to help guide them. The subjects indicated that they relied on the map for information on where to go. The map provided a simple representation of the structure of the text and likely helped subjects determine what information would be coherently related based on the hierarchical structure. There was also some evidence from the protocols of the subjects' awareness of signals in the text, such as titles of the nodes. These too would indicate the relationships between items and would help guide the subjects. Other signals existing in the text may also have been used, such as the frequency of mention and initial sentences, although they were not explicitly stated by subjects. For this hypertext, the map and semantically identifiable labels for the nodes helped the navigation. Without these, subjects may not have had easy ways of gauging where would be the next best place to go in order to maintain coherence.

CHAPTER VI

GENERAL DISCUSSION

Reading Strategies and Comprehension in Hypertext

Despite the differences in the text formats, subjects used strategies to navigate the texts in a coherent manner. The similar approaches to reading the text resulted in similar amounts of text comprehension for the different text formats. For navigating the linear text, subjects merely relied on the linear order of the text for following the coherent path. For the hypertexts, subjects relied on additional constraints in the text such as the map and node titles which signaled the structure of the text. These signals served as guides for locating additional nodes that would flow coherently with the previously read nodes.

Manipulating the goals of the reader changed the readers' strategies. Readers with the specific goals saw fewer nodes and pages of the text and read for a shorter amount of time. However, although they were supposed to find information that was in separate places in the hierarchy, there were no differences in the number of coherent transitions made based on the readers' goals. This indicates that even when readers need to find specific information in an unfamiliar text, they still may need to have the appropriate background context. The verbal protocols from Experiment 2 confirmed that although subjects were aware of where the relevant nodes were, they read down to those nodes in order to put the relevant nodes in the appropriate context. In addition, the goal manipulation did not interact with the text format. Thus, even though subjects had different reading goals, they employed similar reading strategies on the different text formats.

The modified version of the hypertext that provided additional coherence showed no real differences from the original hypertext. Subjects may have ignored the added information that was provided to them by the coherent hypertext. However, based on the strategies used by the subjects in Experiment 1 and on the protocols in Experiment 2, it seems that the general background context provided by the coherent hypertext was the type of information that the subjects needed to maintain coherence and to provide a context for what they were currently reading. Thus, subjects may have not

noticed this information sufficiently or ignored when it appeared. It is also possible that not enough information was provided in these "background context" paragraphs in order to be efficiently used. However, because subjects tended to navigate the coherent hypertext in a coherent manner, this information was also seldom presented to them. Thus, there may not have been enough opportunities for them to use the provided background context to produce significant differences between the hypertexts. Although this study failed to show that the automatic coherence aided the readers, this type of approach may be of use to hypertexts. Given the right type of information in these background context paragraphs and proper training of the readers, readers would not have had to rely as much on following coherent paths. Instead they could used the provided background context to give them the necessary coherence when making jumps out of a particular context.

The strategies used by the subjects show evidence for the dominance of global comprehension in reading. The subjects used a rational approach to reading, maintaining an order of reading that was consistent with the macrostructure of the text. Even when the structure of the text had been modified from its linear form, they chose paths through the text that would flow coherently. This is consistent with the readers' goals to form a coherent macrostructure. Thus, readers of a hypertext are opportunistic. They look for the cues that will lead them to the most coherent path through the text. This behavior is also consistent with rules for coherent reading strategies set out in Goldman and Saul (1990).

With readers using rules to maintain coherence, this means that readers must develop a set of rules and then choose what rules to use for different situations. In this way, reading becomes an application of problem solving. The problem of navigating a hypertext coherently is similar to the problem of search in a problem space in an unfamiliar domain (e.g., Newell & Simon, 1972; Newell, 1980). The reader has a set of possible actions for getting through the text (problem space) and a set of search methods for finding the most effective paths through the text. Since the subjects in this experiment were all expert readers, they had fairly powerful search methods for discovering coherent relationships. However, the subjects were all in an unfamiliar text domain (economics), which put some limits on the ability of the search methods from guaranteeing that they would find the most coherent path.

This approach to navigating through the hypertext is consistent with investigations of exploratory learning in computer systems, such as the CE+ model of Lewis and Polson (1990). Their model of behavior in unfamiliar computer domains uses a label-following hill climbing search method in order to choose the relevant actions. In label-following (e.g., Engelbeck, 1986), the user's goals are matched with the possible actions, and an action is chosen through the degree of overlap of terms between the two. The label following heuristic is similar to the methods the hypertext readers used for

determining where to make the next coherent transition. For finding specific information, subjects could matched their reading goals to the possible nodes they could read to identify relevant nodes of the text and then work their way down to the appropriate nodes. For subjects with general reading goals, they could use the titles as guides for determining the relationships between nodes in order to determine what transitions would be the most coherent.

The Lewis and Polson research suggests that one of the keys to aiding in the problem solving is to design for successful guessing (see also, Norman 1986, 1988). This is a similar necessity for hypertexts. A reader of a hypertext must be able to guess whether following a particular link will both lead toward the relevant information and also be coherent with the current context. In order for a reader to make these successful guesses, this information needs to be evident through the structure of the text and the titles of the nodes.

Text Comprehension Theory and Hypertext

While one of the original goals of this research was to determine the effect of coherence on readers of hypertexts and linear texts, the results suggest that the similar coherence strategies overshadowed the effects of the text formats. In linear text, coherence is a good predictor of comprehension (e.g., van Dijk & Kintsch, 1983; Kintsch & Vipond, 1979). In this research, comprehension was equivalent because the resulting coherence of the text formats were equivalent. Even when modifying the goals of the reader, the text was still read coherently, resulting in few differences in comprehension. Thus, due to a lack of variability in coherence, this research can not draw conclusions about how much the coherence can break down in a hypertext and thereby affect the comprehension.

Readers' strategies indicate the strong role coherence plays in reading a hypertext. In the development of hypertext, the coherence between linked nodes is not often considered. Hypertext guidelines tell creators to link related items, however, there are few guidelines that tell them to make the text of linked nodes cohere. It is typically assumed that because two nodes are linked by some common piece of information, the reader can then generate the correct inferences about the link and incorporate the new information into his or her representation of the text. The results of this study show that subjects avoided the cases of "loosely" linked information (e.g., cross-hierarchical links), and instead primarily made transitions to highly related nodes.

It should be noted that this research examined the strategies for readers who were not familiar with the domain of the text. The results could change greatly with readers who are knowledgeable of the domain. With more background knowledge, readers could perform better problem solving to

choose better coherent paths through the text. These readers would have the correct knowledge to make informed guesses on where to go in the text. With the proper knowledge structures, they could also make non-coherent jumps without adverse affects because they would possess the appropriate background knowledge to make the necessary bridging inferences. Thus, although readers with little background knowledge of the text domain, may have to rely on coherent paths in order to read through the text, a knowledgeable reader may be better able to exploit some of the less coherent links in the hypertext.

Implications for the Design of Hypertexts

The results of this study indicate some issues that should be considered for the design of hypertexts. One of the primary issues is to consider the type of text, how it will be used, and who will be using it. This can affect both what the reader can learn from the text and the type of strategies the reader will employ. In the case of this study, the text was designed primarily for students to study concepts in economics. Students reading the text would have little background knowledge on the topic and would have the goals of either reading through the text or for searching for information in the text. Thus, the results of this study can not be generalized to all hypertexts in all situations.

One of the primary considerations in the design of a hypertext should be to focus on the readers' strategies. These strategies can vary, depending both on the goals of the reader and the background knowledge of the reader. A reader who wants to find specific information in a hypertext needs to have simple methods to locate that information. However having the specific information without the appropriate background context can be useless. Thus, for hypertexts it is not just a matter of getting a reader to the correct place in the text, but also ensuring that the relevant information is represented in a meaningful context. The results from experiment 2 indicated that subjects finding specific information needed the appropriate background context before they could go ahead and read the specific information.

For readers with extensive background knowledge, this meaningful context may not be needed, but for readers without the proper background knowledge, the information will not be as valuable. For this reason, it would make sense to insert the appropriate context, or permit the novice reader to read through all the background context in order to get the relevant information. This concept would be somewhat similar to the idea of "guided tours" in hypertexts in which an author provides a single path through the text with the appropriate context (e.g., Marshall & Irish, 1989). It is also similar to the "training wheels" methodology for interfaces which kept

of contents. This permits readers to get to the relevant places in the text efficiently but also permits them to read the text in the linear order.

One of the current difficulties in developing a hypertext is that hypertext writing environments are still in their infancy (e.g., Britton & Glynn, 1989). Such environments will help a writer see the constraints of the text as the text is developed. This will permit the writer to be better able to consider the coherence of the text from node and how the text should best be structured.

Thus, developing a hypertext is not just a matter of developing a text with all relevant interconnections. Instead, an effective hypertext must be based on a coherent structure. This structure must present an interface to the reader that will make the reader's strategies simple to be carried out.

Future Issues for Text and Hypertext Research

This study shows the importance that readers place on maintaining the coherence of the text as they navigate a hypertext, however there remain many research issues that still must be explored in hypertexts. The results of this study, and many other studies indicate that hypertexts, as they are currently implemented, are not suitable for all applications. Many studies have shown advantages for the linear text or inconclusive differences between the linear text and the hypertext.

While this study sought to identify the effect of differences in the amount of coherence in the hypertext and the linear text, the strategies of the readers created equivalent amounts of coherent transitions. Not all hypertexts would permit subjects to perform as many coherent transitions as the hypertext used in this research. Thus, it would be instrumental to evaluate hypertexts which may cause a greater amount of non-coherent transitions in order to determine how much the coherence affects the comprehension. An alternative is to examine hypertexts that have large amounts of coherence. Hypertexts based on rhetorical models, such as systems for argumentation, enforce the coherence to a large extent. An analysis of reading comprehension on such systems may find that they produce larger amounts of comprehension and also provide a more appropriate representation of arguments than the equivalent linear version.

Modeling Text and Hypertext Readability

A second area research for hypertext is for the development of better models of the text and readability. Current readability measures for linear texts (e.g., Klare, 1964; Peterson, 1954) are not good enough for the evaluation of linear texts, let alone for evaluating hypertexts. These measures typically consider such features as word length and number of syllables, but not such

factors as the coherence of the arguments. As suggested by Kintsch and Vipond (1979), measures of coherence can make accurate predictions of comprehension. However, currently the coherence of the text must be measured through a propositional analysis of the text. This is a time consuming process, involving a thorough hand-crafted analysis of the text. Nevertheless, alternatives to a complete propositional analysis could be performed. One such alternative is to compute the semantic similarity between text sections using semantic models such as latent semantic indexing (Deerwester, Dumais, Furnas, Landauer & Harshman, 1990). These comparisons may then point out hypertext links where the coherence is missing. Thus, modeling the predicted comprehension of a hypertext may not only be useful for research purposes, but also work to provide suggestions of improvements for hypertext developers.

Models of text comprehension (e.g., van Dijk & Kintsch, 1983) can be applied to hypertext as well as to linear text. In both cases the key is to be able to measure the appropriateness of the text's structures and the resulting human's representation of those structures. The primary difference between the current text comprehension models for linear text and those for the hypertext is that a model for hypertext must also take into account the reader's behavior as he or she navigates through the text. In this way, the model must take into account more than just the text, but also the reader's strategies for dealing with the computer interface (e.g., Polson & Lewis, 1990). A model of hypertext comprehension must consider both the information the reader gains from the text and how that information can affect the readers' choice of strategies for proceeding through the text. Thus, this involves both an understanding of the constraints of the text and also of the constraints of the user interface. Given a successful model of these factors, we can generate better guidelines on how to develop more readable hypertexts which support more successful reading strategies.

Where can Hypertext Succeed?

From many of the studies in the field of hypertext, hypertexts have often performed worse, or just as well as the equivalent linear text. This raises the question of what are the areas where a hypertext will be more successful than a linear text, and what sorts of features will the hypertext need for this success.

One of the current successes for hypertext is as a search engine. Although this blurs the distinction between information retrieval systems and hypertexts, the ability to search in through a text simplifies many tasks. One of the remaining problems of paper texts is still the fact that their indexes are inadequate and unwieldy to use. Hypertexts such as SuperBook permit readers to find the relevant information in an efficient manner. Although current computer indexing methods still leave much room for improvement (e.g., Dumais, 1988), the ability to move from one piece of relevant

information to another can not be matched in equivalent paper versions of large amounts of text.

A second area where there is a great deal of potential for hypertext systems is for representing textual information that is not as easily presented in linear form. Such domains as legal argumentation and design knowledge have such problems. Any piece of textual information can have many related arguments and counter-arguments. These can be represented graphically with links between the arguments. Because the arguments are designed to follow each other, they links are constrained to be coherent. Thus the reader of the hypertext may follow the links and still get a coherent representation of the information. While these systems may have an advantage to the reader, they do however, put a large load on the writer. Unlike preparing a single coherent set of arguments, the writer must determine the relationship between all pieces of information and structure them in a much more complicated manner than the linear equivalent. In the future, better computerized methods based on models of the text may be able to do some of this structuring, simplifying the task of the writer.

A third potential area for success in hypertexts is suggested by the investigation in this research of providing dynamic coherence to the hypertext. The goal of dynamic coherence is to provide additional information to the reader so that the text automatically contains information that is more appropriate to the reader's representation of the text. Readers of a text will bring large differences in reading abilities and background knowledge to a text. Through modeling the reader's knowledge, abilities, and goals hypertexts could be developed that do dynamic structuring of the text. These hypertexts could include additional background knowledge to readers with little knowledge and make the text more coherent for readers with low reading abilities. For readers with general reading goals or trying to get an overview, the text could be presented as a single path through the text. For readers trying to find specific information, the text could provide search capabilities and appropriate background context of any relevant items.

While the concept of personalized text sounds like a wonderful idea, accurately gauging the readers background knowledge, abilities and goals can be difficult. A lot of this will depend on being able to develop appropriate user models. These models would need to couple information about what is contained in the text, what information a reader already knows, and what information the reader needs to know.

Success for hypertexts lies in exploiting the powers of both the computer and the writer to generate better personalized texts. However, in order to do this, we must first have accurate models of such factors as: what a reader knows, what a reader needs, what methods a reader can use, and what information is contained in the text. Thus, the future of hypertext will depend on improving both models of the user and models of the text.

ACKNOWLEDGMENTS

I would like to thank my dissertation committee, Walter Kintsch, Peter Polson, Anders Ericsson, Reid Hastie, Gerhard Fischer and Nancy Pennington for their advice and comments on this work. I also thank Adrienne Lee, Evelyn Ferstl, Paula Messamer, Thomas Landauer, Susan Dumais, and Dennis Egan for their help. I would also like to thank the Army Research Institute, the National Institute of Health and Bellcore for providing financial support during parts of this project.

REFERENCES

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

Begoray, J. A. (1990). An introduction to hypermedia issues, systems and application areas. *International Journal of Man-Machine Studies, 33,* 121-147.

Black, J. B., & Bower, G. H. (1979). Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 309-318.

Bolter, J. D. (1991). *Writing Space: The computer, hypertext, and the history of writing*. Hillsdale, NJ: Lawrence Erlbaum.

Botafogo, R. A., & Shneiderman, B. (1991). Identifying aggregates in hypertext structures. In *Proceedings of Hypertext '91,* (pp. 63-74). San Antonio, TX: ACM.

Bransford, J., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717-726.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331-350.

Britton, B. K., & Glynn, S. M. (1989). *Computer Writing Environments: Theory, Research, and Design*. Hillsdale, NJ: Lawrence Erlbaum.

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology, 83,* 329-345.

Britton, B. K., Meyer, B. J., Hodge, M. H., & Glynn, S. M. (1980). Effects of the organization of text on memory: Tests of retrieval and response criterion hypotheses. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 620-629.

Bush, V. (1933). The Inscrutable 'Thirties. *Technology Review, 34*(4), 159-162.

Bush, V. (1945). As we may think. *Atlantic Monthly*(July), 101-108.

Card, S., Moran, T. P., & Newell, A. (1984). *The psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum.

Carroll, J. M., & Carrithers, C. (1984). Training wheels in a user interface. *Communications of the ACM, 27*(8), 800-806.

Charney, D. (1987). Comprehending non-linear text: The role of discourse cues and reading strategies. In *Proceedings of Hypertext '87,* (pp. 109-120). Chapel Hill, NC: ACM.

Conklin, J., & Begeman, M. L. (1989). gIBIS: A tool for all reasons. *Journal of the American Society for Information Science, 40*(3), 200-213.

Croft, W. B., & Turtle, H. (1989). A retrieval model incorporating hypertext links. In *Proceedings of Hypertext '89,* (pp. 213-224). Pittsburgh: ACM.

Crouch, D. B., Crouch, C. J., & Andreas, G. (1989). The use of cluster hierarchies in hypertext information retrieval. In *Proceedings of Hypertext '89,* (pp. 225-238). Pittsburgh: ACM.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391-407.

Doane, S., Kintsch, W., & Polson, P. G. (1990). Action planning: Producing UNIX commands. In *Proceedings of the 11th annual meeting of the Cognitive Science Society, 7* (pp. 30-57). Ann Arbor, Michigan: Erlbaum.

Dumais, S. T. (1988). Textual information retrieval. In M. Helander (Eds.), *Handbook of Human-Computer Interaction.* (pp. 673-700). New York: Elsevier.

Egan, D. E., Remde, J. R., Gomez, L. M., Landauer, T. K., Eberhardt, J., & Lochbaum, C. C. (1989). Formative design-evaluation of 'SuperBook'. *Transactions on Information Systems, 7*(1), 30-57.

Engelbart, D. C. (Ed.). (1963). A conceptual framework for the augmentation of man's intellect. London: Spartan Books.

Engelbart, D. C., & English, W. K. (1968). A research center for augmenting human intellect. In *Proceedings of the 1968 Fall Joint Computer Conference,* (pp. 395-410). Montvale, NJ: AFIPS Press.

Engelbeck, G. E. (1986) *Exceptions to generalizations: Implications for formal models of human-computer interaction.* Unpublished master's thesis, University of Colorado, Department of Psychology, Boulder, Co.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Fincher-Kiefer, R., Post, T. A., Greene, T. R., & Voss, J. F. (1988). On the role of prior knowledge and task demands in the processing of text. *Journal of Memory and Language, 27,* 416-428.

Fischer, G., McCall, R., & Morch, A. (1989). JANUS: Integrating hypertext with a knowledge-based design environment. In *Proceedings of Hypertext '89,* (pp. 105-118). Pittsburgh, PA: ACM.

Fodor, J. D., Fodor, J. A., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry, 6,* 515-531.

Frase, L. T. (Ed.). (1975). Prose processing. New York: Academic Press.

Furuta, R., & Stotts, P. D. (1989). Programming browsing semantics in Trellis. In *Proceedings of Hypertext '89*, (pp. 27-42). Pittsburgh, PA: ACM.

Goldman, S. R., & Saul, E. U. (1990). Flexibility in text processing: A strategy competition model. *Learning and Individual Differences*, 2(2), 181-219.

Gordon, S., Gustavel, J., Moore, J., & Hankey, J. (1988). The effects of hypertext on reader knowledge representation. In *Proceedings of the 32nd Annual meeting of the Human Factors society*, (pp. 296-300). .

Gould, J. D., Alfaro, L., Fonn, R., Haupt, B., Minuto, A., & Salaun, J. (1987). Why reading was slower from CRT displays than from paper. In *Proceedings of the ACM CHI + GI '87*, (pp. 7-11). Toronto, Canada: ACM.

Gould, J. D., & Grischkowsky, N. (1984). Doing the same work with hard copy and with cathode ray tube (CRT) computer terminals. *Human Factors, 26*, 296-300.

Greeno, J. G., & Noreen, D. L. (1974). Time to read semantically related sentences. *Memory and Cognition, 2*, 117-120.

Guthrie, J. T., Britten, T., & Barker, K. G. (1991). Roles of document structure, cognitive strategy, and awareness in searching for information. *Reading Research Quarterly, XXVI*(3), 301-324.

Hammond, N., & Allinson, L. (1989). Extending hypertext for learning: An investigation of access and guidance tools. In *Proceedings of BCS HCI '89*, . Nottingham, UK: .

Hardman, L., & Sharrat, B. (1989). User-Centered hypertext design: The application of HCI design principles and guidelines. In *Proceedings of Hypertext 2*, . York, UK: .

Kieras, D., & Polson, P. G. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 22, 365-394.

Kieras, D. E. (1981). The role of major referents and sentence topic in the construction of passage macrostructure. *Discourse Processes, 4*, 1-15.

Kieras, D. E. (1989). An advanced computerized aid for the writing of comprehensible technical documents. In B. K. Britton & S. M. Glynn (Eds.), *Computer Writing Aids: Theory, Research and Practice*. (pp. 143-168). Hillsdale, NJ: Lawrence Erlbaum.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 363-394.

Kintsch, W., & Keenan, J. M. (1973). Reading rate and retention as a function of the number of propositions in base structure of sentences. *Cognitive Psychology, 5,* 257-274.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilsson (Eds.), *Perspectives on Memory Research.* Hillsdale, NJ: Erlbaum.

Klare, G. R. (1963). *The measurement of readability.* Ames, Iowa: Iowa State University Press.

Landauer, T. K., Egan, D., Remde, J., Lesk, M. J., Lochbaum, C. C., & Ketchum, D. (In press). Enhancing the usability of text through computer delivery and formative evaluation: The SuperBook project. In A. Dillon & C. McKnight (Eds.), *Hypertext: A Psychological Perspective.*

Landow, G. P. (1989). Hypertext in literary education, criticism, and scholarship. *Computers and the Humanities, 23,* 173-198.

Lorch, R. F., & Chen, A. H. (1986). Effects of number signals on reading and recall. *Journal of Educational Psychology, 78,* 263-279.

Mannes, S., & Kintsch, W. (1991). Routine computing tasks: Planning as understanding. *Cognitive Science, 15,* 305-342.

Marchionini, G., & Shneiderman, B. (1987). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer, 21*(1), 70-80.

Marshal, C. C., & Irish, P. M. (1989). Guided tours and on-line presentations: How authors make existing hypertext intelligible for readers. In *Proceedings of Hypertext '89,* (pp. 15-26). Pittsburgh, PA: ACM.

Martin, J. (1990). *Hyperdocuments and How to Create Them.* Englewood Cliffs, NJ: Prentice-Hall.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* Amsterdam: North Holland.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension in ninth-grade students. *Reading Research Quarterly, 16,* 72-103.

Miller, J. R., & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory, 6*(4), 335-354.

Monk, A. F., Walsh, P., & Dix, A. J. (1989). A comparison of hypertext, scrolling, and folding as mechanisms for program browsing. In D. M. J. &. R. Winder (Eds.), *People and Computers IV.* (pp. 421-435). Cambridge: Cambridge University Press.

Nelson, T. (1967). Getting it out of our system. In G. Schechter (Eds.), *Information Retrieval: A critical review.* Wash., D.C.: Thompson Books.

Newell, A. (1980). Reasoning, problem solving and decision processes: The problem space as a fundamental category. In R. Nickerson (Eds.), *Attention and performance VIII.* (pp. 693-718). Hillsdale, NJ: Lawrence Erlbaum.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nielsen, J. (1989). The matters that really matter for hypertext usability. In *Proceedings of Hypertext '89,* (pp. 239-248). Pittsburgh, PA: ACM.

Nielsen, J. (1990). *Hypertext and Hypermedia.* San Diego, CA: Academic Press.

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives in human-computer interaction.* (pp. 31-61). Hillsdale, NJ: Lawrence Erlbaum.

Norman, D. A. (1988). *The psychology of everyday things.* New York: Basic Books.

Nyce, J. M., & Kahn, P. (1991). A machine for the mind: Vannevar Bush's Memex. In J. M. Nyce & P. Kahn (Eds.), *From Memex to Hypertext.* (pp. 39-66). Boston: Academic Press.

Perfetti, C. A., & Goldman, S. R. (1974). Thematization of sentence retrieval. *Journal of Verbal Learning and Verbal Behavior, 13,* 70-79.

Perfetti, C. A., & Roth, S. (1981). Some of the interactive processes in reading and their role in reading skill. In A. Lesgold & C. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ: Lawrence Erlbaum.

Peterson, E. M. (1954). *Aspects of readability in the social studies.* New York: Columbia University.

Polson, P. G., & Lewis, C. (1990). Theory based design for easily learned interfaces. *HCI, 5,* 191-220.

Poulsen, D., Kintsch, E., Kintsch, W., & Premack, D. (1979). Children's comprehension and memory for stories. *Journal of Experimental Child Psychology, 28,* 379-403.

Reder, L. M. (1980). The role of elaboration in the comprehension and retention of prose: A critical review. *Review of Educational Research, 50*(1), 5-53.

Reder, L. M., & Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior, 19,* 121-134.

Samuelson, P., & Nordhaus, W. D. (1989). *Economics* (13 ed.). New York: McGraw Hill.

Shneiderman, B. (1987a). User interface design for the Hyperties electronic encyclopedia. In *Proceedings of ACM Hypertext '87,* (pp. 189-194). Chapel Hill, NC: ACM.

Shneiderman, B. (1987b). User interface design and evaluation for an electronic encyclopedia. In G. Salvendy (Eds.), *Cognitive Engineering in the Design of Human-Computer Interaction and Expert Systems.* (pp. 207-223). New York: Elsevier.

Smith, J. B., & Lansman, M. (1989). A cognitive basis for a computer writing environment. In B. K. Britton & S. M. Glynn (Eds.), *Computer Writing Aids: Theory, Research and Practice.* (pp. 17-56). Hillsdale, NJ: Lawrence Erlbaum.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 275-290.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* New York: Academic Press.

Voss, J. F., Vesonder, G. T., & Spilich, G. J. (1980). Text generation and recall by high-knowledge and low-knowledge subjects. *Journal of Verbal Learning and Verbal Behavior, 19,* 651-667.

Weyer, S. A. (1982). The design of a dynamic book for information search. *International Journal of Man-Machine Studies, 17,* 87-107.

APPENDIX

A. Text from Economics chapter (Samuelson & Nordhaus, 1989)

Every individual endeavors to employ his capital so that its produce may be of greatest value. He generally neither intends to promote the public interest, nor knows how much he is promoting it. He intends only his own security, only his own gain. And he is in this led by an invisible hand to promote an end which was no part of his intention. By pursuing his own interest he frequently promotes that of society more effectually than when he really intends to promote it.

Adam Smith
The Wealth of Nations (1776)

Before we begin to study the economic life of an advanced industrial economy like the United States, we should turn our eyes back to the history and evolution of the modern mixed economy. Centuries ago, government councils or town guilds directed much economic activity in regions of Europe and Asia. However, around the time of the American Revolution, governments began to exercise less and less direct control over prices and economic conditions. Feudal relationships were gradually replaced by what is called the market mechanism, sometimes also labelled "free enterprise" or "competitive capitalism. "

The nineteenth century became the age of **laissez-faire.** This doctrine, which translates as "leave us alone," holds that government should interfere as little as possible in economic affairs. Many governments followed this approach in the middle and late nineteenth century. But before this trend had attained a condition of full laissez-faire, the tide turned the other way. Beginning a century ago, in almost all countries of North America and Europe, the economic functions of government expanded steadily.

Then, around 1980, the tide shifted yet again, as conservative economic policies produced a decline from the high-water mark of intervention in ownership, taxation, and control of the economy.

What are the principles that lie behind the market economy and government's command of economic affairs? In this chapter we will study in detail both these forms of economic organization.

Market, Command, and Mixed Economies

Recall our earlier definitions of market and command economies:

The *market mechanism is* a form of economic organization in which individual consumers and businesses interact through markets to solve the three central problems of economic organization. A *command economy is* one in which the resource allocation is determined by governments, commanding individuals and firms to follow the state's economic plans.

Today, neither of these polar extremes represents the reality of the American economic system. Rather, ours is a *mixed economy,* in which both private and public institutions exercise economic control: the private system through the invisible direction of the market mechanism, the public or government institutions through regulatory commands and fiscal incentives.

One section of this chapter shows how a market mechanism tackles the three problems of economic organization that must be met by any society. The other section briefly reviews the role of the command economy as government directs a modern mixed economy.

HOW MARKETS SOLVE THE BASIC ECONOMIC PROBLEMS

In a country like the United States, most economic questions are resolved through the market, so we begin our systematic study there. Who solves the three basic questions—*how*, *what*, and for *whom*—in a market economy? You may be surprised to learn that no individual or organization is consciously concerned with the triad of economic problems. Rather, millions of businesses and consumers interact through markets to set prices and outputs.

To see how remarkable this fact is, consider the city of New York. Without a constant flow of goods in and out of the city. it would be on the verge of starvation within a week. A wide variety of the right kinds and amounts of food is required. From the surrounding counties, from 50 states, and from the far corners of the world, goods have been traveling for days and months with New York as their destination.

How is it that 10 million people are able to sleep easily at night, without living in mortal terror of a breakdown in the elaborate economic processes upon which the city's existence depends? The surprise is that all these economic activities are undertaken without coercion or centralized direction by anybody.

Everyone in the United States notices how much the government does to control economic activity: it legislates police protection, speed limits, antipollution laws, minimum wages, taxation, national defense, drug prohibitions, highway construction, and so forth. But we often overlook how much of our ordinary economic life proceeds without government intervention.

Thousands of commodities are produced by millions of people, willingly, without central direction or master plan.

Not Chaos but Economic Order

Before they have studied the way the market works, most people see only a jumble of different firms and products. People seldom stop to wonder how it is that food is produced in the right amounts, gets transported to the right place, and arrives in a palatable form at the dinner table. But a close look at the example of New York is convincing proof that a market system is not a system of chaos and anarchy. A market system contains an internal logic. It works.

A market economy is an elaborate mechanism for unconscious coordination of people, activities, and businesses through a system of prices and markets. It is a communication device for pooling the knowledge and actions of millions of diverse individuals. Without a central intelligence, it solves a problem that today's largest computer could not solve, involving millions of unknown variables and relations.

Nobody designed it, and like human society, it is changing. But it does meet the first test of any social organization—it can survive.

History's most dramatic example of the importance of the market economy came in West Germany after World War II. In 1947, production and consumption had dropped to a low level. Neither bombing damage nor postwar reparation payments could account for this breakdown. Paralysis of the market mechanism was clearly to blame. Price controls and overarching government regulation hobbled markets. Money was worthless; factories closed down for lack of materials; trains could not run for lack of coal; coal could not be mined because miners were hungry; miners were hungry because peasants would not sell food for money and no industrial goods were available for them to purchase in return. Markets were not functioning properly. People could not buy what they needed or sell what they produced at free-market prices.

Then in 1948, the government freed prices from controls and introduced a new currency, quickly putting the market mechanism back into effective operation. Very quickly production and consumption soared; once again *what, how,* and *for whom* were being resolved by markets and prices. People called it "an economic miracle," but in fact the recovery was largely the result of a smoothly running market mechanism.

The fact to emphasize is that markets are performing similar miracles around us all the time—if only we look around and alert ourselves to the everyday functioning of the market mechanism. So central are markets to the high levels of output of a capitalist economy that history often witnesses political crises when the market mechanism breaks down. Indeed, a revolutionary out to destroy Western democracies could ask for nothing better than a galloping inflation or depression to paralyze markets and produce political chaos.

THE MARKET MECHANISM

Just how does the market mechanism go about determining wages, prices, outputs, and other economic variables? We can describe its workings easily.

Originally, a market meant a place where goods were bought and sold. Economic histories of the Middle Ages record that market stalls—filled with slabs of butter, pyramids of cheese, wet fish, and heaps of vegetables—would form the commercial centers of villages and towns.1 Today, important markets include the Chicago Board of Trade, where oil, wheat, and other commodities are traded, and the New York Stock Exchange, where titles to ownership of the largest American firms are bought and sold.

Market transactions are often made over the telephone or electronically (as for many financial transactions), and sometimes markets are not centrally organized but operate through individuals buying and selling a house or a car.

A market is an arrangement by which buyers and sellers of a commodity interact to determine its price and quantity.

In a market system, everything has a price—each commodity and each service. Even the different kinds of human labor have prices, namely, wage rates. We receive income for what we sell, and we use this income to buy what we want.

Moreover, prices provide important signals to market participants. If consumers want more of any good—say, wheat—a flood of new orders will be placed for it. As buyers scramble around to buy more wheat, the sellers will raise the price of wheat to ration out a limited supply. And the higher price will encourage greater wheat production.

On the other hand, what if a commodity such as cars becomes overstocked at the going market price? Sellers will lower car prices in their rush to unload unwanted models. At the lower price, more consumers will want cars, and producers will want to produce fewer cars. As a result, a balance (or equilibrium) between buyers and sellers (or what the next chapter and Part Four will call an "equilibrium of supply and demand") will be restored.

What is true of the markets for consumer goods is also true of markets for factors of production. Recall that a factor of production is an input into the productive process—one of the classical triad of land, labor, and capital. If computer programmers rather than typists are needed, job opportunities will be more favorable in the computing field. The price of computer programmers (their hourly wage) will tend to rise, and that of typists will tend to fall. The shift in relative wages will cause a shift of workers into the growing occupation.

Solution to the Three Problems

What happens when we put all the different markets together—wheat, cars, land, labor, capital, and everything else? These form a market mechanism that grinds out an

equilibrium of prices and production. By matching sellers and buyers (supply and demand) in each of these markets, a market economy solves our three problems simultaneously. Here are the bare outlines of such a market equilibrium.

What things will be produced is determined by the dollar votes of consumers—not every 2 or 4 years at the polls, but every day in their decisions to purchase this item instead of that one. The money that they pay into business cash registers ultimately provides the payrolls, rents, and dividends that consumers, as employees, receive as income.

Firms in turn are driven by the desire for profits— profits being net revenues or the difference between total sales and total costs. Firms are lured into production of goods in high demand by the high profits there, leaving behind areas of low profits.

How things are produced is determined by the competition among different producers. The best way for producers to meet price competition and maximize profits is to keep costs at a minimum by adopting the most efficient methods of production. Producers are spurred on by the lure of profit—the production method that is cheapest at any one time will displace a more costly method.

History is filled with examples of how more efficient and lower-cost technologies replaced more expensive ones. Steam engines displaced horses because steam was cheaper per unit of useful work. Diesel and electric locomotives replaced coal-driven ones because of the higher efficiency of the new technologies. In the 1990s, glass fibers and lightwave communications will displace Alexander Graham Bell's traditional copper telephone lines.

We can see the same phenomenon across nations. Bob Smith farms extensively, with much American land relative to each hour of labor. Pierre Reny farms intensively, using much labor to each hectare of French land.

Who makes sure that these *how* decisions reflect the fact that land is scarcer in France than in America? Is it Congress? The National Assembly? The United Nations? Of course not.

The price system is society's signaling device. It tells farmer Smith that he should farm extensively by presenting him with a high ratio of wage rates to land rents. Peasant Reny, on the other hand, faces a low wage/rent ratio and uses more labor per unit of land than does Smith. By looking at price signals, farmers, firms, and other producers can choose the most appropriate technique of production.

For whom things are produced is determined by supply and demand in the markets for factors of production. Factor markets determine the wage rates, land rents, interest rates, and profits—such prices being termed factor *prices*. By adding up all the revenues from factors we can calculate people's incomes. The distribution of income among the population is thus determined by the *amounts* of factors (person-hours, acres, etc.) owned and the *prices* of the factors (wage rates, land rents, etc.).

Be warned, however, that there are also important extra-market influences that affect the resulting distribution of income. This distribution is highly dependent upon the ownership of property, upon acquired or inherited abilities, upon luck, and upon the presence or absence of racial and gender discrimination.

Who Rules?

Who is in charge of a market economy? If we look beyond the details, we see that the economy is ultimately ruled by two monarchs: consumers and technology. Consumers direct by their innate or learned tastes, as expressed in their dollar votes, the ultimate uses to which society's resources are channeled. They pick the point on the production-possibility frontier.

But the available resources place a fundamental constraint on consumers. The economy cannot go outside its PPF. You can fly to London, but there are no flights to Mars. An economy's resources, along with the available science and technology, limit the places where consumers can put their dollar votes.

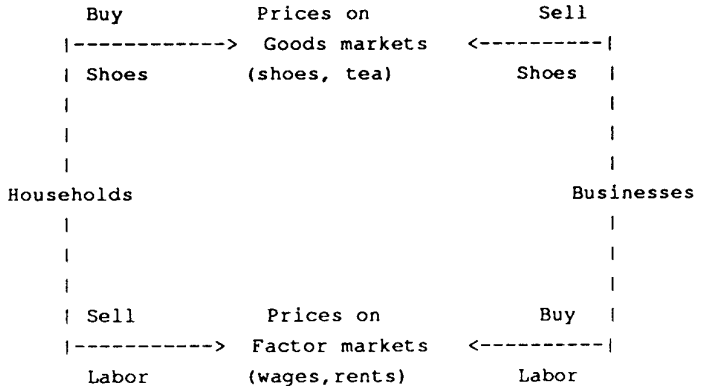
In other words, consumers are not the dictators in deciding *what* goods should be produced. Consumer demand has to dovetail with business supply of goods. Businesses set their prices based on production costs—moving into areas with high profits and leaving unprofitable sectors. So business cost and supply decisions, along with consumer demand, do help to determine *what*. Just as a broker may help arrange a match between buyer and seller, markets act as the go-betweens who reconcile the consumer's tastes with technology's limitations.

It is important to see the role of profits in guiding the market mechanism. Profits provide the rewards and penalties for businesses. Profits lead firms to enter areas where consumers want more goods, to leave areas where consumers want fewer goods, and to use the most efficient (or least costly) techniques of production.

Like a master using carrots and kicks to coax a donkey forward, the market system deals out profits and losses to get *how*, *what*, and for *whom* decided.

A Picture of Prices and Markets

We can picture the circular flow of economic life as shown below. This provides an overview of how market prices reconcile household transactions with business needs. Note two different kinds of markets: one set for outputs like tea and shoes and the second for inputs like land and labor. Further see how decisions are made by two different entities, households and businesses.



Households buy goods and sell factors of production; businesses sell goods and buy factors of production. Households use their income from sale of inputs such as labor and property to buy goods; businesses base their prices of goods on the costs of labor and property. Prices in goods markets are set to balance consumer demand with business supply; prices in factor markets are set to balance household supply with business demands.

All this sounds complicated. But it is just this intricate web of interdependent supplies and demands through which the prices in a market mechanism solve the *how*, *what*, and *for whom*. A few minutes spent studying the above figure now will pay many dividends later in furthering your understanding of the workings of a market economy.

The Invisible Hand and "Perfect Competition"

Adam Smith, whose *The Wealth of Nations (1776)* is the germinal book of modern economics, was thrilled by his recognition of an order in the economic system. Smith proclaimed the principle of the "invisible hand." This principle holds that every individual, in selfishly pursuing only his or her personal good, is led, as if by an invisible hand, to achieve the best good for all. Smith held that, in this best of all possible worlds, any government interference with free competition is almost certain to be injurious. (Reread carefully this chapter's introductory quotation.)

The *invisible-hand* doctrine is a concept for explaining why the outcome of a market mechanism looks so orderly. Smith's insight about the guiding function of the market mechanism has inspired modern economists—both the admirers and the critics of capitalism. After two centuries of experience and thought, however, we now recognize the scope and realistic limitations of this doctrine. We know that the market sometimes lets us down, that there are "market failures," and that markets do not always lead to the most efficient outcome. One of the major market failures, whose consequences will run as a theme through this book, is imperfect competition.

Perfect Competition Smith himself recognized that the virtues of the market mechanism are fully realized only when the checks and balances of perfect competition are present. What is meant by **perfect competition?** It is a technical economic term that refers to a market in which no firm or consumer is large enough to affect the market price. For example, the wheat market is perfectly competitive because the largest wheat farm, producing only a minuscule fraction of the world's wheat, can have no appreciable effect upon the price of wheat.

On the other hand, when a computer company, a steelworkers' labor union, or an airline is large enough to exert an influence on the price of computers, of steelworkers' labor, or of air transportation, some degree of "imperfect competition" has set in.

The invisible-hand doctrine is about economies in which all the markets are perfectly competitive. In such a circumstance, the economy will generate an efficient allocation of resources—where an efficient economy is on its production-possibility frontier. When all industries are subject to the checks and balances of perfect competition, as we will see later in this book, markets will produce the efficient bundle of outputs with the most efficient techniques and using the minimum quantity of inputs.

By contrast, when imperfect competition arises, society may move inside its *PPF*. This would occur, for example, because a single seller (or monopolist) raised the price of a good sky-high in order to earn extra profits. The output of that good would be reduced below the most efficient level, and the efficiency of the economy would thereby suffer. When sellers are few, inadequate checks exist to ensure that prices are determined by the costs of production. And in such a situation, the invisible-hand property of markets may vanish.

In summary, Adam Smith discovered a remarkable property of a competitive market economy: under perfect competition without market failures, markets will squeeze as many useful goods and services out of the available resources as is possible. But where monopolies or pollution or similar market failures become pervasive, the remarkable efficiency properties of the invisible hand may be destroyed.

THE ECONOMIC ROLE OF GOVERNMENT

One section of this chapter described the remarkable efficiency properties of the market mechanism. We saw how an ideal, perfectly competitive economy—where resource-allocation decisions are made through prices and markets—could squeeze the maximum amount of useful goods and services out of the available resources.

But the market does not always behave in an ideal fashion. Indeed, there has probably never been an absolutely pure and perfect competitive market. Rather, market economies suffer

from monopoly and pollution, along with unemployment and inflation, and the income distribution in market economies is sometimes found unacceptable by voters.

In response to these flaws in the market mechanism, democracies have chosen to introduce the visible hand of government into the mixed economy. Governments displace markets by owning and operating certain enterprises (like the military); governments regulate businesses (like telephone companies); governments spend money on space exploration and scientific research; governments tax their citizens and redistribute the proceeds to poor people. In this section we provide a first survey of the rationale and techniques of government intervention in a modern economy.

THE THREE FUNCTIONS OF GOVERNMENT

In discussing government's role, we generally take for granted that government sets the rules of the road, writing laws and enforcing contracts. But what are government's economic functions? They are to promote efficiency, equity, and stability. Government actions concerning *efficiency* are attempts to correct market failures like monopoly. Government programs to promote *equity* use taxes and spending to redistribute income when society shows its concerns for the poor or for particular groups. *Stabilization policy* attempts to shave the peaks and troughs of the business cycle, reducing unemployment and inflation, and promoting economic growth. We will examine briefly each function.

Efficiency

As we saw earlier, economies sometimes suffer from market failures. In actual markets, a firm may profit as much by keeping prices high as by keeping production high. In some markets, firms emit pollution into the air or water or dump toxic wastes into the soil. In each of these cases market failure leads to inefficient production or consumption, and there may be a role for government to cure the disease. But, while evaluating the role of government in curing economic ailments, we must also be alert to "government failures"—situations in which governmental attempts to solve problems may make them worse or cause other problems.

Imperfect Competition One serious deviation from perfect competition comes from *imperfect competition* or *monopoly elements.*

Recall how strict is the economist's definition of a "perfect competitor." The mere presence of a few rivals is not enough for perfect competition. Rather, perfect competition in a market arises when there is a *sufficient number of firms or degree of rivalry such that no one firm can affect the price of that good.* An *imperfect competitor is* one whose actions *can* affect a good's price. In reality, then, almost all business owners, except possibly the millions of farmers who individually produce a negligible fraction of the total crop, are imperfect competitors. At the extreme of imperfect competition is the *monopolist*—a single supplier who determines the price of a particular good by himself.

All economic life is a blend of competitive and monopolistic elements. Imperfect competition, not perfect competition, is the prevailing mode. But to say that a firm can affect the price of its output does not mean it is a dictator. As we shall later see, a business cannot set its prices completely as it pleases and still be profitable. It must take into account the prices of goods that are substitutes for its own. Even if it produces a trademarked heating oil with unique properties, it must reckon with prices charged for other heating oils, as well as for wood, gas, and insulation. Hence, there are always some checks on the economic power of imperfect competitors.

When monopoly power—the ability of a large firm to affect the price in a given market—does become economically significant, we see prices that rise above cost and that depress the amount of output that consumers will buy. The pattern of too high price and too low output is the hallmark of the inefficiencies associated with monopoly power. High prices also

mean high profits, which may be turned to misleading advertising, or even to buying influence from legislatures.

The government does not accept as inevitable all exercise of monopoly power. Since the 1890s, the federal government has imposed both antitrust laws and economic regulation in the name of improving the workings of our imperfectly competitive market system—as Chapters 24 and 25 will discuss.

Externalities A second way in which an unregulated market mechanism may lead to an inefficient outcome arises when there are spillovers or externalities. Look back for a moment to the circular-flow diagram in Figure 3-1. Note that all the transactions between households and businesses take place in voluntary exchange *through markets.* When a firm uses a scarce resource like land, it buys this from the owner in the land market; when a firm produces valuable goods like oil, it receives full value from the buyer in the oil market.

But many interactions in fact take place outside markets. Firm A uses scarce resources like clean air or water without paying those whose air or water is fouled. Firm B, by contrast, decides to provide its employees free vaccinations against a communicable disease; once immunity is achieved, people outside the firm benefit from the reduced dangers of contracting the disease as well. In these cases, an economic bad and an economic good have been transferred outside of market transactions.

Externalities (or spillover effects) occur when firms or people impose costs or benefits on others without those others receiving the proper payment or paying the proper costs.

(You can illustrate a spillover in Figure 3-1 by drawing a direct line from a business to a household. This is an externality because the transaction takes place involuntarily, outside markets.)

As our society has become more densely populated, as production is increasingly based on processes involving harmful substances, negative spillover effects have grown from little nuisances to major threats. Government *regulations* operate with varying degrees of effectiveness to control externalities like air and water pollution, strip mining, hazardous wastes, unsafe drugs and foods, and radioactive materials.

Critics of regulation complain that government economic activity is unnecessarily coercive. Governments are like parents, always saying no: Thou shalt not employ child labor. Thou shalt not pour out smoke from thy factory chimney. Thou shalt not sell dangerous drugs. And so forth.

Many government edicts are in fact controversial: Do we really need to tell firms about how clean they should keep the air inside factories? Do people need to be forced to wear seatbelts?

While the optimal scope of government intervention will never be resolved, most people will agree that government is needed to prevent some of the worst spillovers created by the market mechanism.

Public Goods It is possible to prevent firms from dumping wastes by imposing regulations; it is much more difficult for governments to encourage the production of **public goods.** These are the economic activities—conveying large or small benefits to the community—that cannot efficiently be left to private enterprise. Important examples of production of public goods are the maintenance of national defense and of internal law and order, the building of a highway network, and the support of pure science and public health. Private provision of these public goods will not occur because the benefits of the goods are so widely dispersed across the population that no single firm or consumer has an economic incentive to provide them.

Because private provision of public goods will generally be insufficient, government must step in to provide public goods. In buying public goods like national defense or lighthouses, government is behaving exactly like any other large spender. By casting sufficient dollar votes

in certain directions, it causes resources to flow there. The price system then takes over and ensures that the government-purchased lighthouses or fighter aircraft get produced.

Taxes Government must find the revenues to pay for its public goods and for income-redistribution programs. Such revenues come from taxes levied on incomes, wages, consumer sales, and similar items. Moreover, taxes are raised at all levels of government—city, state, and federal.

Taxes differ from other uses of our incomes in one important respect: Everyone is subject to the tax laws; we are all forced to contribute our share of payments to the government. It is true that the citizenry as a whole imposes that tax burden on itself, and surely we would agree that each citizen has the right to his or her share of the public goods produced by government. However, the close connection between spending and consumption that we see for private goods does not hold for taxes and public goods. I buy a hamburger or a wool sweater only if I want one, but I must pay my share of the taxes used to finance defense, space research, and public education even if I don't care a bit for these activities.

This brief discussion of how government intervenes in markets to improve their efficiency indicates that such actions are firmly grounded in economic logic. Government sets the rules of the road, levies taxes and tolls to pay for collective activities, and buys public goods such as highways, thereby facilitating the smooth driving of private enterprise, preventing abuses when firms become monopolistic road hogs, and curbing firms' activities when their exhaust fumes threaten lives and property.

Equity

Up to now we have focused on defects in the guiding role of the invisible hand—imperfections that perhaps could be corrected by judicious intervention. But assume for the moment that the economy functioned with complete efficiency—always on the production possibility frontier and never inside it, always choosing the right amount of public versus private goods, and so forth. Even if the market system worked as perfectly as just described, many would not consider it ideal. Why not?

In the first place, goods follow dollar votes and not the greatest need. A rich person's cat may receive the milk that a poor child needs to remain healthy. Does this happen because the market is failing? Not at all, for the market mechanism is doing its job—putting goods in the hands of those who can pay the most, of those who have the most money votes. Simply put, even the most efficient market system may generate great inequality.2 If a country spends more on pet food than on paying for college education for the poor, that is a defect of income distribution, not of the market.

Income inequalities may be politically or ethically unacceptable. A nation does not need to accept the outcome of competitive markets as predetermined and immutable; people may examine the distribution of income and decide it is unfair. If a democratic society does not like the distribution of dollar votes under a laissez-faire market system, it can take steps to change the distribution of income.

Often the income distribution in a market system appears to be the result of accidents of technology or birth. Suppose the invention of robots should cause the competitive price of labor to fall greatly, thereby reducing incomes of the poor and turning 95 percent of national income over to robots and their owners. Would everyone regard that as necessarily right or ideal? Probably not. Yet that could be the way the cookie crumbles under the market system. Does someone who inherited 500 square miles of rangeland, for which oil companies offer $50 million per year, necessarily deserve so large an income? People are deeply divided on whether such high incomes should be heavily taxed.

Let's say that voters decide to reduce income inequality. What tools could Congress use? It could engage in *progressive taxation*, taxing a larger fraction of incomes of rich than of poor.

The federal income and inheritance taxes are examples of such redistributive progressive taxation.

Because low tax rates cannot help those who have no incomes at all, governments have in recent decades built up a system of **transfer payments**, or payments to people for which no services are received. Such transfers include aid for the elderly, blind, disabled, and for those with dependent children, as well as unemployment insurance for the jobless. This system of transfer payments provides a "safety net" to protect the unfortunate from economic destruction. Governments sometimes subsidize consumption of low-income groups by providing food stamps, subsidized medical care, and low-cost housing.

Through the process of economic growth and welfare programs that established minimum standards of living, much of the great and visible destitution of nineteenth-century capitalism has been erased in the twentieth century. But relative poverty and deprivation have proven difficult to eradicate.

What can economics contribute to debates about equality? Economics as a science cannot answer such normative questions as how much of the competitively determined incomes—if any—should be transferred to poor families. But it can analyze the economic costs or benefits of different redistributive systems. Economists have devoted much time to analyzing whether different income-redistribution devices (such as taxes and food stamps) lead to social waste (such as people working less or buying less nutritious food). They have studied whether giving poor people cash rather than goods is likely to be a more or less efficient use of society's resources. Economics is like a good travel agent. You, the traveler, must decide whether you want sun or snow. But once you choose, the agent can help you get to your destination quickly and cheaply.

Stability

In addition to its role as promoter of efficiency and equity, government also engages in the macroeconomic function of promoting economic stability.

Since its origins, capitalism has been plagued by periodic bouts of inflation (rising prices) and depression (very high unemployment). Sometimes these episodes were so violent, as in the German hyperinflation during the 1920s, that social turmoil, revolution, and war followed in their wake.

At times, as during the Great Depression in America in the 1930s, hardship persisted for a decade because political leaders had insufficient economic understanding to take steps for economic revival. More recently, in the early 1980s, governments here and abroad took steps to reduce high rates of inflation; these measures led to high unemployment and declining inflation.

Today, thanks to the intellectual contribution of John Maynard Keynes, his followers, and also his critics, we have a better understanding of how to control the ups and downs of the business cycle. We now understand that, by careful use of the government's monetary and fiscal powers, the levels of output, employment, and inflation can be influenced. The *fiscal powers* of government are those just discussed— the power to tax and to spend. The *monetary powers* involve regulating banks and the financial system with an eye toward determining the supply of money, interest rates, and credit conditions. Through these two central tools of macroeconomic policy, governments can influence the rate of growth and level of output, the level of employment, and the price level of an economy.

Governments in advanced industrial countries successfully applied the lessons of the Keynesian revolution over the last half-century. Spurred on by expansionary monetary and fiscal policies, the market economies of the world witnessed a period of unprecedented economic growth from World War II until the early 1970s. In the 1970s, however, the advanced economies encountered foul economic weather: two major oil-price increases, harvest failures,

commodity shortages, a breakdown in the international financial system, and an increasingly heavy burden of government regulatory and redistributive programs. As economic growth slackened and inflation soared, people became skeptical about the ability of monetary and fiscal policies to attain the goals of macroeconomic stabilization.

Today, policymakers are realizing that a modern economy faces a fundamental macroeconomic dilemma: *No country has for long periods of time succeeded in having free enterprise, low inflation, and full employment.* Just as today's market economy cannot have both guns *and* butter, so a market economy cannot simultaneously attain full employment *and* stable prices.

The circular flow diagram of prices and markets presents a succinct summary of the economic role played by government today, showing why, as government actively promotes efficiency, equity, and stability, the United States in the 1980s is called a "mixed economy." Indeed, in all advanced industrial economies, the market determines most individual prices and quantities while government steers the overall economy with programs of taxation, spending, and regulation. Both halves— market and government—are essential for a soundly functioning economy. To operate a modern economy without both is like trying to clap with one hand.

B. Questions used in the posttest

1. What is a "Perfect Competitor"?

Consider the following cases of government intervention in the economy. For each one, circle what role of government is being pursued in each case.

2. Pollution limitations
   a) Efficiency
   b) Equity
   c) Stabilization Policy

3. The food-stamp program
   a) Efficiency
   b) Equity
   c) Stabilization Policy

4. Price regulation of AT&T or other large companies
   a) Efficiency
   b) Equity
   c) Stabilization Policy

5. A monetary policy step to curb inflation
   a) Efficiency
   b) Equity
   c) Stabilization Policy

6. What is a mixed economy?

6a How does a mixed economy differ from a command and a market economy?

7. Who decides what things should be produced in a market economy?

8. What is the "Invisible Hand"?

9. Who developed the concept of the "Invisible hand"?
   a) Karl Marx
   b) Adam Smith
   c) John Maynard Keynes
   d) Thomas Payne

10. This chapter had two main subtopics. They were:
(circle two)
    a) Market, command, and mixed economy
    b) How markets solve the basic economic problems
    c) The solution to 3 problems
    d) The economic role of government
    e) The invisible hand and "Perfect Competition"

11. What two methods can the government use to bring about equity?
(circle two)
    a) Subsidize consumption
    b) Regulate externalities
    c) Progressive taxation
    d) Monetary policies
    e) Stabilization policies

12. Which of the following titles was a subtopic of the section called "The market mechanism"
    a) Public goods
    b) Not chaos but economic order
    c) The three functions of government
    d) The solution to 3 problems
    e) Stability

13. How easy was the format of the economics text to read?

Very easy Very difficult

1----------2----------3----------4----------5----------6----------7

(circle a number)

14. How easy was the text itself to understand?

Very easy Very difficult

1----------2----------3----------4----------5----------6----------7

(circle a number)

15. How easy was it to find specific pieces of information in the text?

Very easy Very difficult

1----------2----------3----------4----------5----------6----------7

(circle a number)

16. Please estimate how long the whole economics text was in terms of the number of words and pages.

I think this text was _____ words long

I think this text was _____ pages long

17. How many years have you used computers for such things as word-processing or graphics?

_____ Years.

18. How many years have you used computers for programming?

_____Years.

19. How many years have you used Macintosh computers?

_____Years.

University of Colorado at Boulder

Institute of Cognitive Science

Muenzinger Psychology Building
Campus Box 344
Boulder, Colorado 80309-0344
(303) 492-5063
FAX: (303) 492-7177
email: cogsci@psych.colorado.edu
URL: http://psychwww.colorado.edu/ics/