Comprehension and Learning Through Multimedia: Integrative Processing of Text and Illustrations

Marguerite Roy
Department of Educational and Counselling Psychology
McGill University, Montréal

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Educational Psychology, Specialization in Applied Cognitive Science

0-612-70141-7

Acknowledgments

I wish to extend my heart-felt appreciation to Dr. Carl H. Frederiksen for the excellent guidance and advice I received from him over the years. His knowledge and enthusiasm for conducting research in the field are truly inspirational. I am particularly grateful to him for encouraging me to participate in so many different aspects of research and teaching.

I would also like to say thank you to my committee members, Dr. Alenoush Saroyan and Dr. Alan Breuleux, for the their support, advice, and patience they provided.

I am also indebted to the support and of Dr. Suzanne Lajoie, Dr. Janet Donin, Dr. Robert Bracewell, and Dr. Michael Hoover for their overall support and encouragement.

It was my good fortune and pleasure to work along side so many great colleagues and friends at the lab. I would like to acknowledge the support of Nancy, Sonja (is it time for lunch yet?), Tina, Claudia, Barbara, Roger, Lorraine, Mauro, Gloria, and Hélène. Thank guys for all the stimulating talks.

I would also like to acknowledge the contribution of André Renaud for his help in creating the computer display environment and his encouragement. Thanks also go out to Julien Mercier who kindly provided the translation of the abstract.

Extra special thanks go out to Dave, my partner in life, for believing in me and helping me to become unstuck in thesis time. I also wish to acknowledge the kind support of Nancy and Owen Buckingham who provided a warm and friendly source of moral support.

I would also like to extend my appreciation to the staff and students in the Mechanical Engineering Department at McGill University. The enthusiasm and interest they demonstrated was truly appreciated.

Finally, I would like to acknowledge the support of FCAR (grant # 921807) and McGill Social Sciences Research Grants for supporting this research.

Abstract

The comprehension of illustrations and text was studied from a cognitive discourse processing perspective. Typically, learners must construct conceptual knowledge representations that integrate different types of information from diverse sources and modalities (e.g., text and illustrations). Currently, little is known about how such integrative processing works in "multimedia" learning situations. This study focused on the semantic representations that low prior knowledge learners constructed as they read from text and static illustrations presenting multiple types of information (structural, functional, and energy) describing a functional system. Both the text and the illustrations were modified so that structural information would be highlighted over other information types.

Participants were twenty-four undergraduate engineering students who had little prior knowledge of the target domain (the human visual system), but were experienced in learning about functional systems using texts and illustrations. Six students were randomly assigned to each of four presentation conditions: (a) text only, (b) illustrations only, (c) text with controlled access to illustrations, or (d) text with free access to illustrations.

The materials were presented individually in a computer environment which recorded and timed all information accessed. Participants provided on-line interpretations as they read, post-input verbal and visual free recalls of the materials, and responses to integrative comprehension questions. Planned comparisons were used to contrast: (a) the two text with illustrations groups, (b) the combined text with illustrations groups to the text only group, (c) the text with illustrations groups to the illustrations only group, and (d) the processing of information which was privileged (structure) to other information.

The results indicated that the text and illustrations each provided mutually constraining information that functioned together to support comprehension. Illustrations aided the construction and elaboration of mental models by providing an external context that supported more active conceptual processing and integration of information. Text aided both literal and high level comprehension by communicating the meaning of illustrations and signaling what information was important. These results support perspectives on situated learning which emphasize the role that discourse plays in comprehending knowledge in environments involving more than one external information source. Future research is recommended to extend such findings to other populations of learners and materials.

Résumé

La compréhension d'illustrations et de texte a été étudiée dans une perspective de traitement cognitif du discours. Typiquement, les apprenants doivent construire des représentations conceptuelles qui intègrent différents types d'information provenant de diverses sources et modalités. L'état actuel de la recherche ne renseigne pas sur la mécanique d'un tel processus de traitement intégratif dans un contexte d'apprentissage "multimédia". La présente étude porte sur les représentations sémantiques construites par des apprenants possédant peu de connaissances antérieures lorsqu'ils lisent un texte et des illustrations statiques. Ce matériel présente plusieurs types d'information (structurel, fonctionnel et concernant l'énergie) qui décrivent un système fonctionnel. Le texte et les illustrations ont été modifiées afin que l'information structurelle soit mise en évidence par rapport aux autres types d'information.

Les participants sont vingt-quatre étudiants au baccalauréat en ingénierie possédant peu de connaissances antérieures sur le sujet du matériel (le système oculaire humain), mais expérimentés dans l'apprentissage de systèmes fonctionnels par des textes et des illustrations. Six étudiants ont été assignés, au hasard, à l'une des quatre conditions de présentation du matériel : (a) texte seulement, (b) illustrations seulement, (c) texte avec accès contrôlé aux illustrations, ou (d) texte avec accès libre aux illustrations.

Le matériel est présenté individuellement dans un environnement informatique qui garde la trace de l'information accédée ainsi que de la durée de cet accès. Les participants produisent des interprétations en temps réel lorsqu'ils lisent, des rappels libres verbaux et visuels après la lecture et des réponses à des questions de compréhension et d'intégration. Des comparaisons planifiées sont utilisées pour comparer: (a) les deux groupes avec texte et illustrations, (b) les groupes avec texte et illustrations combinées avec le groupe avec texte seulement, (c) les groupes avec texte et illustrations avec le groupe avec illustrations seulement, et (d) le traitement de l'information privilégiée (structurelle) avec les autres types d'information.

Les résultats indiquent que le texte et les illustrations produisent des informations qui se contraignent mutuellement et qui fonctionnent de concert dans la facilitation de la compréhension. Les illustrations favorisent la construction et l'élaboration de modèles mentaux en fournissant un contexte externe qui permet un traitement conceptuel plus actif ainsi que l'intégration de l'information. Le texte favorise tant la compréhension littérale que la compréhension de haut niveau en communiquant la signification des illustrations et en signalant l'information importante. Ces résultats supportent les perpectives concernant l'apprentissage contextualisé, lesquelles mettent l'accent sur le rôle que le discours joue dans la compréhension des connaissances dans des

environnements impliquant plus d'une source externe d'informations. Des recherches futures sont recommandées afin de bonifier ces résultats auprès d'autres populations d'apprenants ainsi qu'avec du matériel différent.

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

Text and Illustrations as Information Sources

Information presented in text and graphic form differ qualitatively. Written text presents conceptual and relational information as a verbally encoded linear sequence of propositions and is an effective means for describing abstract and general information. In contrast to written texts (or verbally encoded information in general), graphics and pictorial representations present information simultaneously to the learner, and tend to be well suited for displaying spatial, concrete, and specific information. Such external representations convey information by the way in which their components are spatially organized (Thorndyke & Stasz, 1980) and thereby require use of the specific capabilities of the human visual system for the perception of spatial configuration (Schnotz, 1993a). Comprehension of illustrated information, however, also requires a level of conceptual analysis whereby properties of the visual stimulus configuration are interpreted and mapped onto certain semantic structures (Pinker, 1990). That is, just as for text, graphic information must be represented and interpreted semantically if it is to be understood (Palmer, 1975; 1978; Pylyshyn, 1973).

It has been suggested that in certain learning situations the differences between encoding and presenting information in textual versus graphical form may offer certain advantages to the learner (e.g., Kosslyn & Pomerantz, 1977). For example, when trying to plot a route for driving to a new destination, consulting a graphical representation such as a map, is generally more helpful than consulting a textual description of the geographical layout of towns, cities, highways, etc. On the other hand, abstract information is not easily represented graphically. So for example, when trying to understand an abstract concept, a textual description is generally preferred over an illustration. Commonly, however, a combination of text and graphic representations are used to convey information. The question then arises as to how the learner uses both modes of information in conjunction to understand the encoded conceptual and relational information.

Although there is general agreement that information presented in textual and graphical form differ qualitatively, it is also accepted that both sources of information must be represented and interpreted within the confines of the human cognitive system. However, we still do not know precisely how readers are able to integrate text and graphic information. The purpose of the present study is to gain a better understanding of this process.

The Problem: Cognitive Processing of Illustrated Text

Text is usually one of several sources of information available for learning in a domain. Commonly, other external representations (e.g., illustrations) are also used and provide their own ways of representing information. In addition, computer based multimedia learning environments which combine text, graphic, sound, and animation are becoming more prevalently used as commonly available sources of information. In such situations, a learner must be able to construct a conceptual knowledge representation (i.e., mental model) that integrates different kinds of information from diverse sources and modalities. Currently, however, we know little about how integrative processing works in such "multimedia" learning situations.

A large part of the reason why little in known about this issue is that the study of comprehending and learning in multimedia environments is a complex research problem that involves an integrated set of assumptions concerning the content and structure of information in each medium, the learner, the task, and the learning context. There is a clear need to bring well defined models and methods for studying comprehension and learning processes to the study of learning in multimedia environments. Furthermore, as several researchers have pointed out the (Arens, Hovy, & Vossers, 1993; Collins, 1996; Saloman, 1989) the design issues involved in studying the effectiveness of various learning environments are vast and involve many potential cost-benefit relationship that are as yet unclear.

The present study is designed to investigate the issue of how readers use and understand illustrations and text from a cognitive discourse processing perspective. An attempt is made to address such broad questions as: How do readers make use of information presented through a combination of text and illustrations? How do characteristics of text influence the reader's use of illustration information? How do characteristics of illustrations affect the reader's processing of text? How is overall comprehension influenced by such processes? Are multimedia environments (such as text and illustrations) advantageous for students' comprehension and learning? Or do they impose serious problems for students due to increased processing load and demands for integration of information within and across media? Surprisingly, there is little research which adequately addresses these basic questions.

Why Study Comprehension and Learning From Illustrated Text?

There are several practical reasons to justify why it is important to understand how people comprehend and learn from illustrated text. Perhaps the most obvious reason is that illustrated text makes up a substantial portion of academic and instructional materials across domains and across levels of education. In fact one team of researchers (Stone & Glock, 1981) have commented on the prevalence of illustrated text by suggesting that study of the reading process itself should properly include the study of illustrated text. Illustrated texts have a long tradition of instructional use. Gombrich (1990) provided a historical review of some of the uses of pictorial instruction. While he described some examples of wordless pictorial instructions (e.g., airline instructions for what to do in case of an emergency) he reported that, more commonly, language is used in conjunction with pictures to communicate a message. Some examples of the range of domains in which one can find historical use of illustrated text include: religion (particularly in the middle ages); botany (e.g., herbals for identifying and learning about plants); medicine (e.g. medical treatises describing various internal and external parts of the body) ; ornithology (e.g., manuals for identifying and learning about birds); the military arts (e.g., manuals describing military maneuvers); dance (e.g., manuals describing dance steps); geography (e.g., maps); mining (e.g., describing the overall setup and steps involved in mining various materials); the visual arts (e.g., manuals describing the steps involved in drawing particular objects); sports (e.g., manuals describing how to perform certain sports including fencing, wrestling, and gymnastics); and various forms of handiwork (e.g., Origami, Needlework, cotton weaving, and tying various knots).

The omnipresence of textual illustrations continues today. Illustrations are currently used in a variety of instructional materials ranging from primers to college textbooks to technical manuals (Kozma, 1991). Given the prevalence of illustrated texts, it may be surprising to note that the incorporation of various graphic devices in instructional text has traditionally been based on the intuition of writers and text designers rather than on research findings (Guri-Rozenblit, 1988). Clearly, a better understanding of how people comprehend and learn from illustrated text would have important ramifications for the design of instructional texts in general.

A second and related reason to be concerned with this issue is grounded in the observation that while a great deal of instructional time and effort is devoted to teaching students how to read and comprehend text, there is virtually no emphasis placed on attempting to teach students how to interpret and comprehend illustrations as information sources. This instructional gap may stem from the assumption that illustrations are self explanatory. However, as will become apparent, this assumption is unfounded. Exactly how to proceed to remedy this situation is not at all obvious given that our current understanding of what constitutes "graphic literacy" is poor. A better understanding of the cognitive requirements and relevant dimensions involved in constructing meaning from illustrations could be extremely helpful in generating instructional goals and methods for teaching students how to interpret and integrate information from text and illustrations.

Yet another reason for studying how readers comprehend and learn from illustrated text is that research findings could be used to develop more general theories of comprehending and learning from multiple sources of information. For example, theories and research findings on illustrated texts could potentially be extended to research concerned with learning from multimedia. Presently, the development of multimedia learning environments abound despite the fact that we have a very limited understanding of how people comprehend, integrate, and learn from multiple media sources (e.g., Rieber, 1990a; 1990b; 1991a; 1991b). The design and instructional efficacy of multimedia environments and our ability to assess their effectiveness may depend fundamentally on our understanding of the cognitive demands for successfully learning in such environments, within which texts and illustrations play a major role.

Finally, the issue of how people are able to learn from and integrate information across multiple sources and media is relevant to contemporary theories of situated cognition. Such theories (e.g., Brown, Collins, & Duguid, 1989; Greeno, 1989; 1991; Resnick, 1987; Suchman, 1988) have focused on the issue of how features of a particular learning or performance situation may help to contextualize and thereby constrain cognitive performance. For example, Greeno (1989) describes a theoretical model in which knowledge is defined in terms of the relationship between the individual and the social or physical situation. In particular, such theories have extended the notion of "learning situation" to include multiple participants, multiple representations, and multiple media. While such a position extends the traditional information processing paradigm from the study of internal structure and process, to the consideration of both internal and external structures and processes, clarification of how such factors combine to influence cognition and performance have yet to be clearly specified.

On the other hand, attempting to adopt a situated cognition perspective to the study of illustrated text can be helpful in that it re-focuses the problem from one of trying to determine the conditions under which text illustrations influence learning and comprehension to consideration of how text and illustrations may serve as mutually constraining sources of information for the reader. It is within this context that the question of knowledge integration becomes particularly relevant. From this perspective, learning from illustrated text may be viewed as a somewhat more constrained learning task than learning from text alone. That is, one of the functions that text illustrations may serve is precisely to help the reader *con*textualize text based information. Thus, a better understanding of the how individuals learn from illustrated texts may yield important information that is relevant to more clearly specifying models of situated learning in general in that such findings may help to delineate/describe how learners use and comprehend information from more than one source.

Chapter II: Literature Review

At the most basic level this study is concerned with investigating the cognitive processes that readers use to construct an integrated conceptual understanding from information expressed through text and illustrations. As such, several areas of literature are pertinent to this study including theories and empirical research on (a) reading illustrated text, (b) the comprehension of illustrations, (c) text comprehension in general, and (d) mental models in particular.

The literature review is organized into several main sections. The first section provides a brief description of early research (up to early 1970s) on the effects of illustrated texts. This research demonstrates mixed results, but is fraught with methodological problems.

The second section concentrates on research dating from between the mid 1970s to the mid 1980s. This work demonstrates general support for the notion that illustrations aid comprehension and memory of text. Two major criticisms of this work, however, are that it approaches the problem of comprehending and learning from illustrated text in an a-theoretical manner, and secondly, that it tends to focus on the products of comprehension and learning rather than on the processes. As such, this research is not capable of describing *how* illustrations affect comprehension and learning from text.

The third section reviews the literature on the cognitive processing of illustrations in an attempt to obtain some indication of how illustrations are interpreted by learners and what cognitive processes are affected. One shortcoming of this research, however, is that it provides an incomplete picture of the problem as it does not address the issue of how illustrations are interpreted and understood in the context of text.

The next main section focuses on research which has attempted to study patterns of individual and group differences in the use and effectiveness of text illustrations in promoting comprehension. This research is important in addressing the issue of how learner characteristics and behavior influence the effectiveness of text illustrations. Again, however, this research can be criticized on the basis that it has tended to focus on the products rather than the processes of comprehension.

The next section reviews research which demonstrates that text can affect the processing of illustration information. Such findings are significant, particularly when combined with the findings that illustrations facilitate text processing, in emphasizing the reciprocal nature of comprehension effects that can occur between media. It is in this context that the issue of knowledge integration, both within and across media, becomes relevant. This in turn leads to the following section reviewed which provides an overview of the concepts of coherence, cohesion, and knowledge integration as they have been studied in the area of text comprehension.

The seventh section reviewed explores current models of discourse processing in an attempt to bring well developed theoretical models and methods to the study of comprehending and learning from illustrated text. It is concluded that adopting such a perspective greatly clarifies the problem of how to study the cognitive aspects of understanding and learning form illustrated text. The final section focuses on a review of research concerned with investigating the role of situation or mental models in learning from text in general and from illustrated text in particular.

Early Research on Learning From Illustrated Text

While the literature is replete with studies examining the effects of text illustrations on learning, there is relatively little work conducted prior to 1970. Samuals (1970) reviewed some of this early research from which he concluded that text illustrations are *not* generally beneficial, and in some cases may pose a hindrance to learning (e.g., Samuals, 1967; Weintraub, 1960; Willows, 1978; Vernon, 1953). However, at the time of this review, there were very few studies investigating the effects of text illustrations on the learning of content. That is, most of these studies were concerned with investigating the use of illustrations as facilitators of *learning to read* rather than *reading to learn*. Furthermore, no distinction had yet been made between *text relevant* and *text irrelevant* illustrations (an issue which will be elaborated below). In addition, many of these studies tested the ability of children to read words in isolation and thus, it could be argued that such a decontextualized task misrepresents the potential effects of illustration on learning that occurs in more natural situations. Yet another shortcoming of Samuals review is that it gave preference to studies in which pictures were used as adjuncts to text rather than as additional sources of information in their own right. These shortcomings are probably responsible for the mixed results observed (see Bluth, 1981 ; and Schallert, 1980 for an expanded review and discussion of the shortcoming of these early research efforts).

Product Oriented Research on Learning From Illustrated Text

Following this period, investigations into the psychological effects of illustrated text on comprehension and learning were largely evaluative in nature. The research question focused upon at this time was whether or not adding illustrations to text yielded any facilitory or inhibitory effects on learning and comprehension of text content. However, the general research paradigm used tended to focus on assessing the products of comprehension and learning rather than the processes.

For example, Goldberg (1974) studied the effects of adding illustrations to study materials used in a fifth grade spelling and grammar lesson. Half of the students were provided with the "normal" study sentences and half were additionally provided with illustrations that depicted the content of the study sentences. The results demonstrated that students who received the sentences with illustrations performed significantly better on multiple choice and fill in the blank questions. Using a similar design, Holliday (1975) compared the performance of tenth graders on a multiple choice test for an extended text on biology and found significant improvement for students who also received labeled line drawings.

Rasco, Tennyson, and Boutwell (1975) demonstrated that the positive effects of illustrations on readers' ability to recall information described by both mediums extended to adults as well. These researchers performed a series of experiments to investigate the effects of imagery and illustrations on learning from text. In the first study, college students read an extended text about revolutions. Half of the students were instructed to form mental images of the concepts as the read and half were not. In addition, half of each of these groups received illustrations to help them understand the text. Students who received either imagery instructions, actual illustrations, or both were better able to answer questions about the text than students who received no imagery instructions or illustrations. There was no difference between the imagery, illustration, or combined groups. In their second and third studies, Rasco et al. used the same procedures and design but with fifth and sixth graders, and fourth and fifth graders respectively. Similar results were reported.

Anglin (1986) also reported two experiments demonstrating positive effects of pictures on adult readers' memory for information described by both text and illustrations. Graduate students who read text with illustrations remembered more of the information described by both media than subjects' who read the same passage without illustrations. This effect was equally strong following a 14 day delay, and a 26 day delay.

Sherman (1976) investigated the effects of text illustrations on high school students' ability to comprehend abstract and concrete passages. In addition, Sherman studied the effects of illustration placement (i.e., before or after the text) on comprehension performance. Sherman found that, overall, illustrations significantly facilitated students' ability to recall main idea units. However, the facilitory effect of illustrations was strongest when placed at the beginning of an abstract passage.

Bock and his colleagues (Bock & Hormann, 1974; Bock & Milz, 1977) provided an early demonstration that illustrations can facilitate sentence processing. In these experiments, people's memory for noun (e.g., "The man washed his car.") and pronoun sentences (e.g., "He washed it.") were compared under three presentation conditions. In the control condition, subjects were simply presented with the sentences. In a second condition, subjects were simultaneously presented with line drawings depicting the content of the sentence. In a third condition, subjects were presented with the drawings and sentences sequentially rather than simultaneously. The results indicated that, when presented together, illustrations led to increased retention for pronoun, but not noun sentences. However, when presented sequentially, illustrations were found to facilitate memory for both pronoun and noun sentences. Thus, it appears that the effects of illustrations on text processing vary with the temporal locus of illustration consultation. It should be pointed out, however, that this study was limited to investigating how a single illustration affects the processing of a single sentence. Thus, it is not clear how such results generalize to reading situations which involve the processing of connected discourse accompanied by multiple illustrations.

In their now classic study, Bransford and Johnson (1972) demonstrated that subjects were better able to recall an obscure passage and rated the text as more comprehensible when they were presented with a relevant picture which depicted the components of the paragraph in a meaningful configuration prior to hearing the text. However, no facilitation occurred for preceding pictures which depicted the same objects, but whose spatial configuration did not represent the context of the text. Nor, did facilitation occur when pictures appeared at the end of the text. The authors concluded that pictures can function like headings, thereby activating relevant schemas and influencing the way in which subsequent text will be processed.

Some researchers have suggested that texts with pictures may be more enjoyable for readers thus inducing them to invest more cognitive effort in processing (Williams, 1968). According to this hypothesis, pictures should facilitate performance in a general way, and not just for content that is illustrated. Furthermore, there should be no systematic relationship between text and illustration characteristics and comprehension outcomes.

Levin's (1981b; 1989) work (discussed in more detail below) provides strong contradictory evidence for this position. In particular, Levin's findings on the importance of illustration relevance (i.e., decorative illustrations have not been shown to provide any comprehension benefits) argues forcefully that what information gets depicted in an illustration does indeed have important implications for readers' comprehension and memory. This is not to say that illustrations have no motivational effect on readers, rather the point is that the motivation hypothesis in and of itself is too simplistic to account for the data in general.

Other researchers have suggested that illustrated texts (in particular text relevant pictures) may serve as a source of explicit repetition of information, thus allowing the reader an additional processing opportunity and facilitating reader comprehension and memory. Theorists who hold this view argue that there is nothing inherently special about representing information graphically. According to this hypothesis, pictures should only facilitate comprehension and memory of information that is presented in both textual and illustrated form.

For example, Daneman and Ellis (1995) recently conducted a study in which they compared the relative effects of pictorial and non pictorial supplements in promoting memory for text. Undergraduate students were assigned to one of three presentation conditions: a text with no pictorial or verbal aids condition; a text with pictures condition (simple line drawings depicting either an individual idea or the relation between two or more ideas); and a text with verbal captions condition. The results indicated equal facilitation effects on subjects' free recalls and responses to multiple choice questions for the pictorial and verbal aids conditions. These findings were consistent across three expository texts that varied in the amount of spatial content they conveyed. The authors concluded that, for the kinds of materials they studied and when text and illustrations are equated in terms of information content, illustrations are no better than verbal captions at enhancing memory.

However, there is also evidence for the contrary position. For example, Levin, Bender, and Lesgold (1976) explicitly tested the repetition hypothesis by comparing a reading condition in which each text sentence was orally presented twice to conditions in which text and illustrations, or text only were presented once. While this repetition of information improved children's memory for text, recall was still significantly higher for the text plus illustration condition.

Related to the repetition hypothesis, is the issue of whether or not illustrations can facilitate the processing of non illustrated text information. The data regarding this question is also somewhat confusing. That is, some studies have found support for the position that illustrations can facilitate the processing of non illustrated text information (e.g., Royer & Cable, 1976; Sherman, 1976; Small, Lovett, & Scher, 1993; Weisberg, 1970) while others have not (e.g., Haring & Fry, 1979; Peeck, 1980; Rusted & Coltheart, 1979; Small, Dowling, & Lovett, 1981).

It sum, a variety of theoretical explanations as to how illustrations affect learning have been offered including, attentional, affective, and cognitive. Potential cognitive effects that have been identified include information repetition, schema effects, concretization of information, among others. However, it is not at all clear how one is to distinguish a particular explanation as more favorable given that this research has not adequately measured cognitive processing per se.

Based on such research, one is left to conclude that illustrations may serve a variety of functions. This observation led to the development of the functional approach to the study of illustrated text.

A functional approach to the study of illustrated text. It has been pointed out that some researchers have approached the study of learning from illustrated text as if all illustrations can be expected to produce similar effects. However, "not all text illustrations are created equally", nor are they intended to serve identical functions. Several researchers have attempted to clarify the importance of making this distinction by creating taxonomies designed to classify text illustrations in terms of their cognitive functions. (e.g., Duchastel & Waller, 1979; Dwyer, 1988; Hegarty, 1992; Levin, 1981b; Levin, Bender, & Pressley, 1979; Levin & Berry, 1980; Mayer, 1993a; Shriberg, Levin, McCormick, & Pressley, 1982).

For example, Duchastel and Waller (1979) have claimed that having a set of principles relating illustration characteristics to potential effects on the reader would be a first step towards sorting out the literature on learning from illustrated text. Duchastel (1978) has described three main roles that illustrations in text may serve: (a) attentional, (b) explicative, and (c) retentional. Duchastel and Waller (1979) note that instructional texts make primary use of explicative illustrations which are used to directly teach, explain, and clarify ideas. These authors go on to identify seven non-mutually exclusive functions that explicative illustrations may serve, including: (a) *descriptive* (i.e., displaying what an object looks like), (b) *expressive* or emotive, (c) *constructional*, where the intent is to explain how various component parts of an object fit together, (d) *functional* which enable the learner to visually follow through the unfolding of a process or the organization of a system, (e) *logico-mathematical*, (f) *algorithmic* which display action possibilities, and (g) *data display* which enable quick visual comparison and easy access to data.

Levin (1981a), who is perhaps the most widely cited source on this issue, has proposed five different functions that text illustrations may serve. Levin distinguishes between illustrations which serve: (a) decoration , (b) representation , (c) organization , (d) interpretation , and (e) transformation functions. Essentially, decorative illustrations refer to illustrations that are generally used to increase the attractiveness of a text. As such, decorative illustrations may serve two functions: to increase publisher sales, and to increase the readers motivation. In either case, these illustrations do not directly relate to text content and are therefore not expected to directly affect cognitive processing or to otherwise facilitate learning. Representational illustrations generally add concreteness to text by *retelling* content in pictorial form. Organizational illustrations function to add coherence to text content by providing an organizational scheme. Illustrations with an interpretive function can aid prose learning by adding comprehensibility to hard-to-understand content. Such illustrations are often used as advanced organizers to describe unfamiliar or difficult concepts. Transformational illustrations are designed to serve a mnemonic objective by helping the reader to recode, relate, and retrieve text content.

Moderate to substantial learning benefits have generally been observed when text illustrations are *relevant* (i.e., overlap or are redundant with) to the text content (e.g., Haring & Fry, 1979; Levin & Lesgold, 1978; Levie & Lentz, 1982; Readence & Moore, 1981; Schallert, 1980), while illustrations that serve a decoration function (i.e., do not overlap, or conflict with text content) are not beneficial to learning from text (e.g., Levie & Lentz, 1982; Levin, Anglin, & Carney, 1987).

Mayer (1993a) describes a somewhat modified taxonomy of text illustrations. He identifies four different kinds of text illustrations on the basis of the cognitive processes affected. The first kind of illustrations he identifies is the *decorative* illustration. Decorative illustrations are pictures that fill space on the page without enhancing the text message and thus serve no cognitive function. *Representational* illustrations may be used to depict a single element described in the text and affect the readers ability to select (devote attention to) certain text information. *Organizational* illustrations depict the elements of an object and the relationship between elements. This type of illustration affects the readers ability to select and organize text information. Finally, *explanative* illustrations are used to explain how a system works. Explanative illustrations affect the readers ability to select, organize, and integrate text information with their prior knowledge.

Criticism of the functional approach. Weidenmann (1989) has criticized the functional approach as being too restrictive. He points out that the functional approach has failed to adequately study the range of possible interactions between reader, illustration, text, and task characteristics. Instead, Weidenmann calls for a more cognitively oriented approach whereby researchers extend their focus on the function of illustration types to the analysis of *how* illustrations are used by various readers in information processing terms, for various texts, and for various tasks. For example, asking the questions "How does the reader interpret the task?", "How does this interpretation influence processing of text and illustrations?", "How do readers use illustrations?", and "How do text characteristics influence readers use of illustrations?" may provide more insight into the mechanisms underlying the effects of illustrated text on comprehension and learning. Indeed many researchers have levied similar criticisms against the product-oriented approach in general to the study of illustrated text (e.g., Baggett, 1989; Mayer, 1993b).

Summary of product oriented work. Research between the period of the mid 1970s to the mid 1980s used a general paradigm comparing an experimental condition in which an illustration was added to a text with a control condition which included a text only version. Employing product measures of comprehension and learning, these studies have demonstrated strong support for the notion that adding *relevant* illustrations to texts improves performance on recognition, cued, and free recall measures. This effect has proven to be highly robust and general across age, text types and levels of education (e.g., Alesandrini, 1984; Anglin, 1986; Duchastel & Waller, 1979; Levin & Lesgold, 1978; Levie & Lentz, 1982; Schallert, 1980).

While these studies have demonstrated a positive effect for adding illustrations to text, it is difficult to make any specific conclusions about how illustrations affect comprehension and learning. A major short coming of this work is that it does not attempt to describe which aspects of comprehension are specifically affected by the addition of illustrations. Research on how readers understand and learn from illustrated text should be motivated by and grounded in a well developed theory of comprehension processing. Another problem that such studies face is that they do not provide an adequate description of the content and structure of text or illustrations, nor do they adequately address how such factors may interact with characteristics of the reader to affect comprehension.

In sum, while such product oriented research has been pivotal in answering the question of whether or not illustrations can facilitate learning, it has been unable to answer the question of how illustrations facilitate learning. In order to begin to answer this question, the literature on the cognitive processing of illustrations is reviewed next.

Cognitive Processing of Pictorial Information

Several researchers have attempted to describe the various kinds of information that can be represented and understood through illustrations. For example, Mandler and Johnson (1976) and Mandler and Parker (1976) have investigated the role of organization on people's memory for various kinds of information in complex pictures. Mandler and Johnson (1976) presented subjects with pictures which were either *organized* (i.e., the pictorial elements were arranged so that they described a coherent scene) or *unorganized* (i.e., the pictures contained elements of an organized scene, but were incoherently arranged). The picture information was analyzed into four different types: (a) inventory information, which specifies what objects are depicted, (b) spatial location information, which specifies the relative location of depicted objects, (c) descriptive information, which specifies the figurative detail of depicted objects, and (d) spatial composition information, which specifies figure/ground information. Subjects were exposed to illustrations for a moderate duration (between 5 and 60 seconds) and the accuracy and latency of subjects' recognition memory for each type of picture information was tested using various foils. The results indicated that memory for inventory and descriptive information did not vary across the condition of picture organization. Thus, the processing of this information was unaffected by organization. However, spatial location information was better recognized for organized pictures, whereas, spatial composition information was better recognized for unorganized pictures. Mandler and Johnson concluded that organization (i.e., coherence) specifically affects processing of spatial information in illustrations.

Using similar pictorial stimuli, Mandler and Parker (1976) followed up this work by testing undergraduate's recognition and recall memory for these different information types either immediately or following a one week delay. Again, organization was found to have a large effect on memory for spatial location information, both immediately and following the delay. In contrast, descriptive information appeared to be unaffected by organization. Thus, it appears that spatial information is processed somewhat differently than other sorts of illustrated information.

More recently, Bieger and Glock (1985; 1986) have extended Mandler's taxonomy of picture information for the purpose of analyzing picture-text instructions. In addition to the categories of pictorial information described by Mandler and her colleagues, Bieger and Glock identified the following kinds of semantic information: (a) operational information, which directs an implied agent to engage in a specified action, (b) orientation information, which describes the spatial orientation of an object, (c) contextual information, which provides the theme or organization for other information that may precede of follow it, (d) covariant information, which specifies a relationship between two or more pieces of information that vary together, (e) temporal information, which provides information about the time course of events or states, (f) qualifying information, which modifies other information by specifying the manner, attributes, or limits of that information, and (g) emphatic information, that directs attention to other information.

These categories of information were derived from a problem solving study in which Bieger and Glock had a group of subjects think aloud as they performed assembly tasks. Subjects were provided with completely assembled objects to guide their performance. The think aloud data were analyzed using discourse analysis techniques to identify the various kinds of semantic information that people used in performing the assembly tasks. Based on these data, Bieger and Glock hypothesized that inventory, operational, spatial, and contextual information were critical sources of information for the successful completion of an assembly task. They tested this hypothesis in a second study by comparing the assembly performance of subjects who received complete (i.e., inventory, operational, spatial, and contextual information) or incomplete instructions that were presented through either text, pictures, or a combination of text and pictures. The results revealed that readers who received complete information demonstrated faster assembly times with fewer errors than subjects who received incomplete information. Thus, each of these types of information contributed significantly to subjects' task performance. Unfortunately, in this study, Bieger and Glock did not explicitly compare the effectiveness of various modes of information presentation.

However, in a more recent study, Bieger and Glock (1986) did compare the effectiveness of mode of presentation of instructions on subjects' ability to successfully complete assembly tasks. Using stimuli similar to that used in their previous study, Bieger and Glock varied the mode (i.e., text or illustration) in which they presented subjects with operational, spatial, and/or contextual information. Their data revealed that textual presentation of spatial information slightly reduced the number of assembly errors that subjects made, while pictorial presentation of this information led to very large decreases in assembly times. Pictorial presentation of contextual information slightly reduced the number of assembly errors and assembly times. No differences were found for the presentation of operational information.

Based on the above studies, it appears that spatial configuration information presented through illustrations is processed differently than other sorts of pictorial information (i.e., inventory, descriptive, spatial composition). In particular, it seems that the effective comprehension of configuration information is affected by coherence relations (i.e., organization). Furthermore, with regard to conveying procedural information, there is some evidence to suggest that illustrations may be more efficient at communicating spatial information than text.

It should be noted, however, that the research literature also reports evidence that our representation of visually presented spatial information is influenced and organized by non spatial knowledge (e.g., Stevens & Coupe, 1978; McNamara, 1986; Tversky, 1981). For example, Stevens and Coupe (1978) performed a series of experiments in which subjects studied simple maps and then were asked to judge the relative position of two real or fictitious cities from memory. Of particular interest, was the study of the errors resulting from inferential processes in spatial judgments. Superordinate structure (e.g., Nevada is east of California) systematically distorted how spatial relationships were remembered by subjects (e.g., San Diego (California) is north-northwest of Reno(Nevada)). Based on the pattern of results, Stevens and Coupe concluded that spatial information is stored hierarchically in the form of a semantic network representation where spatial relationships between regions that are part of the same superordinate region are stored explicitly, but relations between regions of different superordinate regions must be inferred.

There are also studies which have demonstrated significant semantic effects on memory for pictorial information. For example, Goldstein and Chance (1971) report data showing that people's recognition memory for complex pictures with little meaningful information (e.g., snowflakes) is much poorer than their ability to remember equally complex pictures but with highly meaningful information (e.g., faces).

The point to be made is that spatial information may be organized by non spatial information. Such finding demonstrate that the processing and use of illustrated information is not limited to perceptual analysis, but rather includes semantic interpretation and integration with other kinds of knowledge.

Processing of domain-specific illustrations. Several researchers have concerned themselves with investigating how individuals process illustrations that are conventionally used to communicate information in specific domains. In general, this research has been important in providing evidence that the comprehension of illustrations is not necessarily an "automatic" "obvious" or "effortless" process. Rather, this research suggests that the effective interpretation of illustration information in specific domains can be characterized as a kind of expertise that is associated with specific knowledge attained through prolonged periods of experience and practice with such materials.

For example, DeGroot (1965) provided an early demonstration that people's memory for meaningful chess board configurations was far superior than for random arrangements for subjects who were experienced chess players. However, no such differences were found for inexperienced players. The results were interpreted as demonstrating that one of the key features that differentiates more from less experienced individuals in a particular domain is the ability to efficiently interpret and recall information that is specifically relevant to the domain in question.

Lowe (1993) conducted a study in which he used drawing tasks to examine differences between the ability of experienced meteorologists (experts) and non meteorologists (novices) to comprehend and recall weather maps. He found both quantitative and qualitative differences in how experts and novices copied and recalled map information. Specifically, Lowe found that the meteorological experts demonstrated superior recall and tended to reproduce map information in terms of domain relevant high-level abstract relations, while non-meteorlolgists appeared to focus on low level visuo-spatial characteristics. This pattern of results is in agreement with classic expert-novice differences which demonstrate that experts tend to organize information in terms of domain principles, whereas novices tend to rely on surface features (e.g., Chi, Glaser, & Farr, 1988).

Lesgold, Rubinson, Feltovich, Glaser, Klopfer, and Wang (1988) studied the performance of individuals who varied in their radiological diagnosis expertise (residents and senior staff) on their ability to diagnose chest X-rays. Lesgold *et al.* also report differences that are consistent with those classically found in studies comparing expert and novice performance. Of specific relevance to the current topic were the findings that experts spent proportionally more time generating a problem representation than novices and were more likely to use domain specific language when describing the X-rays, while novices were more likely to use spatial terms.

In a problem solving study, Frederiksen, Bédard, and Roy (in preparation) examined the ability of experienced engineers and less experienced engineering students to solve shear force and bending moment problems. Solving problems in this particular domain requires that the performer generate several different kinds of external representations describing the forces assumed to be operating on a static object. These representations include: (a) a free-body diagram which represents a schematized illustration of the problem situation, (b) a set of equilibrium equations, which mathematically describe the magnitude of forces operating for a given location, (c) a shear force diagram, displaying the magnitude and location of vertical forces operating at various points along the object, and (d) a bending moment diagram, indicating the location and magnitude of twisting forces operating along the object. As part of this study, subjects were asked to explain how each of the representations produced during problem solving related to one another, and how they represented the principles of physics that were assumed to be operating. The results of this task revealed strong differences between more and less experienced subjects. In particular, the data demonstrated that highly experienced performers were much more adept at explaining specifically what information each representation provided, how each representation reflected relevant principles of physics, and how these representations were related to one another in domain-specific terms.

The findings of the above studies provide evidence that interpretation of domain specific graphic representations is associated with specific domain knowledge and experience. However, there is also evidence that domain experience in and of itself maybe insufficient to distinguish more and less successful performance on the processing of conventionalized graphic representations.

For example, Gobert (1989) employed methods from the cognitive analysis of text comprehension and problem solving to study the nature of semantic representations involved in the interpretation of graphic information sources used in the domain of architecture. A group of architectural experts and sub-experts were asked to think aloud as they studied the architectural plans of an actual building. In addition, both the expert and sub-expert groups varied in the degree to which they had specific prior knowledge of the building represented in the plans. Gobert employed the methods of propositional and frame analyses to extract the semantic content of subjects' verbal protocols. She coded protocols for the following kinds of semantic information: *single* object descriptions, including identity and class relations, attribute relations, and description of geometrical form; and *relative* object description, including descriptions of function, relative location, circulation and access, composition, and structure. No significant overall differences in the semantic content of the protocols were found between the expert and sub-expert group. However, subjects who had *specific* prior knowledge of the building, regardless of whether they were classified as experts or sub-experts, provided both a greater amount and a different pattern of object descriptions than subjects who did not have any specific prior knowledge of the building. Gobert's study is noteworthy on two counts. First, her findings indicate that *specific* prior knowledge exerts powerful effects on the processing of graphic information. Second, her study exemplifies how the methods of text comprehension can be extended to the study of graphic comprehension processing.

In another study, Thorndyke and Stasz (1980) studied individual and group differences between expert and novice map users on their ability to learn from maps. Over the course of several trials, subjects were required to think aloud as they studied two different maps. After subjects had "learned" the material, they were required to reproduce the maps and to answer several open-ended integration and inference questions. The think aloud protocols were analyzed to identify the processes that subjects used. The data yielded an array of individual differences in the pattern of procedures that individuals used to learn the relevant information. These differences, however, were found to distinguish "good" from "poor" learners rather than the individual's level of experience with map use. Thus, neither the highly experienced nor the low level experienced subjects were uniform in their processing and ability to learn from maps. Procedures that distinguished more and less successful learners were related to differences in attentional, encoding, evaluation, and control processes. Such results suggest that something more than domain general experience is operating when a graphic representation is being processed.

Illustrations and problem solving. In addition to evidence that illustrations can facilitate memory for prose information, there is also research which supports the notion that illustrations can facilitate problem solving performance in ways that are different from text. Although this literature is not specific to the issue of learning from illustrated text, it is relevant to the question of what cognitive processes may be facilitated by illustrations. Both the effects of using pre-constructed illustrations and the effects of generating illustrations on problem solving performance are briefly reviewed below.

In an early study, Stone and Glock (1981) found that college students who read an illustrated text describing a procedure for assembling a model of a loading cart made significantly fewer errors in applying the procedure than students who received an informationally equivalent text-only version. The illustrations used in this study were simple line drawings depicting the cart at various stages of assembly. The illustrations were designed to be completely redundant with the text in terms of content. Thus, the illustrations themselves added no new content, but served to make the information more concrete in the sense described by Levin (1981a).

Larkin & Simon (1987) have investigated the computational efficiency of verbal versus diagrammatic external representations for solving problems in the domains of mathematics and physics. They note that one major difference between these two forms of external representation is that information is organized sequentially in text, but spatially in diagrams. The authors reasoned that, when the two representations are informationally equivalent, any observed performance differences must be due to differences in the computational efficiency of the information processing operators that problem solvers use to operate on them. In their investigations, Larkin and Simon did indeed find that the two forms of presentation differed in the capabilities they support in terms of problem solvers' ability to recognize patterns, carry out direct inferences, and their use of control strategies (in particular, search).

Larkin and Simon suggest some of the reasons why "good" diagrams might be superior to verbal descriptions for solving problems. They note that diagrams can group information that is used together, thus allowing the solver to avoid large amounts of search for the elements needed to make a problem-solving inference. In addition, diagrams typically use location to group information about a single element, thus allowing the solver to avoid the need to match (i.e., integrate) related symbolic labels. Finally, diagrams can support a large number of perceptual inferences, which tend to be extremely easy for humans to make. Thus, Larkin and Simon argue that diagrams can be "better" representations, not because they present more or different information, but because the indexing of this information can be used to support useful and efficient computations. They warn, however, that even "good" diagrams are only useful to those who know the appropriate computational processes for taking advantage of them.

In her own work, Larkin (1989) has described a model of display-based problem solving to explain both why certain tasks are made easy and why certain mistakes are common in working with external graphical displays. She argues that use of displays may considerably reduce the complexity of the mental processes involved in solving a problem. Larkin notes that an "appropriate" diagram can substitute easy perceptual judgments for more error-prone and effortful logical judgments. However, use of diagrams may also lead to errors due to the invisibility of function and the non-salience of form. Diagrams which effectively support problem solving are capable of representing all information which is relevant to the problem solution in a well constructed form. Thus, some of the characteristics of "good" diagrams would include the qualities of informational completeness, appropriate organization, and transparency.

In a recent study, Zhang (1987) studied the effects of different isomorphic versions of Tic-Tac-Toe on undergraduates problem-solving. His results demonstrated regardless of whether certain aspects of the way the task was externally represented were consistent, inconsistent, or irrelevant to the task such features none the less affected performance. That is, the form of the external representation determined to some extent what information was perceived, what processes were activated, and what structures were discovered by participants. Thus, his results demonstrate that both perceptual and cognitive biases are operating when working with external representations.

Bauer and Johnson-Laird (1993) have also reported similar results to those of Larkin and Simon (1987), although they approach this area from a mental model point of view. They found that undergraduates were more effective and efficient at performing a deductive reasoning task involving double disjunction when the premises were presented diagrammatically rather than verbally. The authors argue that diagrams can facilitate the ability to construct mental model representations of the problem. Specifically, diagrams facilitate mental model construction and problem solving by (a) making problem relevant information explicit, thereby enabling the solver to keep track of alternative possibilities and to discount inconsistent conclusions, and (b) allowing the solver to directly construct mental models (i.e., the diagrams explicitly depict the problem situation) and bypass the process of constructing meaning from verbal premises, thereby freeing up working memory.

Stenning and Oberlander (1994) have made similar claims with regard to the ability of Euler circles to assist in reasoning tasks. They argue that such diagrammatic representations can assist logical reasoning by making some information transparent thereby limiting the need for abstraction.

Hong and O'Neil (1992) studied variation among instructional strategies used to help learners build relevant mental models in the domain of inferential statistics. Two main instructional factors were studied: presentation sequence of declarative and procedural information (separate with conceptual preceding procedural instruction vs. combined), and presentation mode (diagrammatic vs. descriptive). Graduate and undergraduate students were assigned to one of the four instructional conditions. Learning was assessed through the analysis of learners' think aloud protocols while solving a series of problems following instruction. The results revealed a significant instructional advantage when conceptual information was described prior to procedural information, and when extensive use of diagrams was made. There was no significant interaction between the two factors. One of the main conclusions that Hong and O'Neil made was that the use of diagrams significantly aided learners in their ability to adequately represent the structure of problems and their underlying concepts.

In addition to investigations of the effects of presenting problem solvers with pre-constructed illustrations, there has also been some research on the question of how generating graphical displays can facilitate problem solving performance. For example, Cox and Brna (1995) studied how students generate and use external graphic representations to solve analytic problems. To investigate the spectrum of external representations that students naturally use while solving such problems, these researchers collected the sketches that students generated while working on Graduate Record Exams. These data demonstrated that students use a broad range of external representations to support their problem solving including: plan diagrams, tabular representations, directed graphs, set diagrams, logic, lists, and natural language. In addition, Cox and Brna collected data using a computer-based system (switchER) which was constructed to dynamically record subjects' behavior as they constructed and reasoned about various external representations. SwitchER was developed in order to study the process and time-course of the manner in which people use external representations while solving problems. Their results indicate that subjects' prior knowledge of the formalisms of external representations was an important predictor of their use of such representations. Thus, this particular study provides additional support for the notion that specific knowledge and experience may underlie peoples ability to effectively use graphic devices.

Summary of processing of illustration information. In general, research concerned with investigating the cognitive processing of illustrations has demonstrated the special status of spatial information. Specifically, it appears as though spatial information is processed differently than other sorts of graphically depicted information. The literature investigating the effects of illustrations on problem solving performance suggests that organizing information spatially in external graphics can have positive effects on problem solving performance.

Research has also demonstrated that specific knowledge and experience contribute significantly to the effective interpretation of domain specific illustrations. However, it must be stated that such studies focus on areas in which there are strong conventions for interpreting and using illustrations. Twyman (1985) has noted that, in such cases, pictures tend to carry the status of sublanguages and involve precise rules which are often understood only by specialists. As such, the above studies represent specialized uses of illustrations rather than more common everyday uses.

With regard to research on the effects of illustrations on problem solving performance, both the use of pre-constructed displays and the act of generating visual displays of information can aid problem solving. Specifically, the cognitive benefits of external graphic representations on problem solving performance can include a reduction in search and memory load. Since graphic representations organize information by location, they can facilitate problem solving by supporting perceptual judgments thereby allowing the solver to bypass search, comprehension, and inference processes that are more heavily demanded by linguistically encoded information (Bauer & Johnson-Laird, 1993; Larkin & Simon, 1987). It has also been suggested that because graphics present concrete and definite representations of objects, they serve to represent the structure of a problem in a constrained way (Larkin, 1989). This quality of "weak expressiveness" of graphics can facilitate solving performance and make inferences more tractable provided that the displays are well constructed, are capable of representing all information relevant to the problem, and the solver knows how to take advantage of such representational qualities.

It should also be noted, however, that the facilitory effects of graphics on problem solving are not necessarily equivalent across tasks or domains. For example, with regard to the comprehension of computer programs, there is some evidence that text representations tend to be more effective at supporting problem solving than graphic representations (Petre & Green, 1992). Furthermore, while there is support for the notion that illustrations may be particularly effective at conveying spatial configuration information, it is not yet clear what other sorts of information may be effectively communicated through depiction. Nor is it clear how illustrated information may effect processing of other kinds of textually presented information, or how text information may influence the processing of illustration information.

According to the research reported above, to some extent illustrations are processed differently than texts. It appears as though illustrations in general are particularly effective at conveying spatial information and at organizing information spatially. Such differences can have facilitative effects on memory, the generation of inferences, as well as on problem solving performance. However, with regard to the processing of domain specific illustrations, it appears as though the performer's ability to reap such benefits depends on his/her level of specific prior knowledge and experience with using such materials. In addition, there is some data to support the idea that there are additional factors that contribute to individual differences in terms of how readers process illustrated information.

While this research is helpful in terms of exploring what cognitive processes may be facilitated by illustrations, it does not address the question of how illustrations are interpreted and integrated with text. The next section reviews in more detail the role that individual characteristics may play in the effective use and comprehension of illustrated texts.

Individual and Group Differences in Readers' Use of Text Illustrations

There is a body of evidence which indicates that comprehension benefits due to adding illustrations to text vary with reader characteristics. For example, Rankin and Culhance (1969) used a cloze task to assess graduate and sixth grade students comprehension for illustrated and unillustrated text material. These researchers found that text illustrations resulted in comprehension benefits for graduate students, but not for sixth graders. In a more recent replication of this study was conducted by Reid, Briggs, and Beveridge (1983) with 14-year-olds. These researchers also failed to find a facilitory effect of illustrations suggesting that in contrast to older (and presumably more able) readers, younger readers do not necessarily benefit from text illustrations when comprehension is measured literally.

Reid and Beveridge (1986) investigated the relationship between picture facilitation, children's ability, and text difficulty. A group of 13- to 14-year-old students who were following the same science curriculum were ranked in terms of their science ability on the basis of a common within school examination and divided into four ability groups designated as "superior", "above average", "below average", and "inferior". Students read expository text in one of the following four conditions: text of easy readability with pictures, text of easy readability without pictures, text of difficult readability with pictures, and text of difficult readability without pictures. In addition, half of these groups read traditional print materials, while the remaining half were presented with materials on a computer screen. Learning was assessed using a set of post-reading multiple choice questions. Questions focused on information that could be answered from the pictures only, the text only, and through information presented through both text and pictures. The results indicated no effects for text difficulty, and no overall effects for the addition of picture in general (i.e., information available in pictures but unavailable from the text). However, the addition of specific pictures (i.e., pictures which were redundant with text information) was found to facilitate learning for the "superior" and "above average" ability groups, but to interfere with learning for the "below average" and "inferior" ability groups. Additionally, there was some evidence that learning through traditional print materials was superior to learning from computer displayed information. The authors argue that these results support the notion that picture facilitation effects depend on the reader's ability in the target domain.

More recently, Reid and Beveridge (1990) attempted to extend research on how children benefit from illustrated text to more closely examining how children who differ in their abilities to learn from illustrated text may be using different text processing strategies. Building on previous research which demonstrates that children of various abilities may differ in the extent to which they benefit from text illustrations (e.g., Beveridge & Griffiths,1987) Reid and Beverage (1990) investigated the possibility that children of different abilities may employ different strategies when reading illustrated text.

Reid and Beverage had a group of 14-year-old students read three brief illustrated expository science texts on a computer. The three texts were selected to vary according to difficulty. A single illustration was used for each text and was designed to overlap with the main conceptual content (i.e., the main idea) of text content. The computer was configured to allow readers access to either text or illustration information from any point. The text was displayed a sentence at a time and accumulated on the computer screen as the reader progressed through each sentence. The computer was programmed to provide a trace of when and where information was accessed by the reader. This experimental setup enabled collection of the following on-line data: (a) reading time for each sentence in the text, (b) access time for the illustration, (c) frequency of illustration access, and (d) position of illustration access in relation to the last sentence read. These data provided the evidence from which reading processes and strategies were inferred.

Three criterion-referenced tests were developed to assess learning. Questions were designed so that they could be answered correctly from information available in either the text or the illustration. Performance on these tasks was used to rank order subjects into six equal groups, from those learning least to those learning most. The most difficult text was dropped from the analysis because it was too difficult for the children. The results indicated that: (a) illustrations were looked at significantly longer as text difficulty increased; (b) the amount of learning decreased as the time spent looking at illustrations increased; (c) the ratio of illustration inspection time to text inspection time was different for children who differed in their ability to learn from the text. The least successful children spent four seconds consulting the text for every second they spent inspecting the illustration, while the most successful children spent six seconds consulting the text for every second spent inspecting the illustration; and (d) the least successful children accessed the illustrations significantly more frequently than the most successful children. Thus, the least successful children focused on the illustrations both more often and for proportionally more time than they focused on text information.

Comprehension differences were not due to a difference in text processing time, as the lower achieving children did not differ from their more successful counterparts in this regard. Furthermore, it is clear that the poorer performance of these children was not due to a failure to use illustrations. The authors suggest that one possible interpretation of these findings is that low ability students shift their attention from the text to illustration at inappropriate points in text processing thus disrupting their ability to construct a coherent text representation. While this interpretation is interesting, Reid and Beveridge do not provide any data or analysis to back up this possibility.

There are additional problems with the methodology and the interpretation offered by Reid and Beveridge. First, it is apparent that while the authors are concerned with investigating and making conclusions about text processing strategies they do not offer any a priori expectation or explanation of what possible strategies children might use to process and learn from illustrated text. As such, their explanations for their findings appear post hoc. Nor did they attempt to measure or control readers' prior knowledge. Also, from the manner in which the experimenters aggregated children into groups of more and less successful learners it is not known whether these children differed in their general text processing abilities or whether they differed only in their ability to learn from illustrated text. Furthermore, the data collected to index processing strategies simply provides on-line information about the pattern of text-illustration consultation but needs to be supplemented with information about what children are doing with this information as they proceed through the materials.

In a more recent study, Schnotz, Picard, and Hron (1993) used think aloud protocols to compare the processing of successful and unsuccessful learners as they read a text about time and place along with a map showing the time zones of the world in a hypertext environment. These researchers found that successful learners used both map and text information differently than unsuccessful learners. In particular, successful learners appeared to use map information more intensively than less successful learners. In addition, although the two groups did not differ in their ability to recall the text, more successful learners were more likely to recall information relevant to a mental model of the text.

In a similar vane, Oshima, Scardamalia, and Bereiter (1996) recently investigated differences between a group of grade 5-6 students who achieved more and less conceptual understanding in the domain of electricity. One of the major differences noted between these two groups was that the more successful students were more likely to make use of graphic representations than their less successful counterparts.

More recently, Davidson-Shivers, Shorter, and Jordan (1999) employed the think aloud technique to study the learning strategies and navigational decisions that fifth grade students made while learning a Hypermedia lesson. These researchers found that the group of students who scored highest on a post-test had demonstrated a greater variety of learning strategies on-line than did students who scored in either in the average or low post-test score ranges. In addition, while the low ability students did show evidence of on-line constructive processing, their protocols tended to contain conceptual errors.

Thus, the literature relating illustration use to learner success is somewhat inconsistent. Some studies report that successful learners make more use of illustrations (e.g., Oshima, Scardamalia, & Bereiter, 1996; Schnotz, Picard, & Hron, 1993), whereas other studies report that less successful learners rely more on illustration information (Reid & Beveridge, 1990). Additional research is needed to investigate this issue further. It is likely, however, that such differences are more appropriately explained by variation in *how* readers are processing text and graphic information in an ongoing manner and whether or not they possess adequate prior knowledge rather than by differences in intensity or frequency of use. For example, Kuntz, Drewniak, and Schott (1992) demonstrated that readers with low prior knowledge are more likely to benefit from illustrations. Kuntz et al. (1992) studied how high and low prior knowledge university students studied text alone or text with one of two supplements (either a representational picture or tree diagram). For high prior knowledge readers, the presence of pictures did not facilitate comprehension and the presence of a tree diagram actually interfered with comprehension. However, low prior knowledge readers appeared to benefit from either supplement.

Further evidence that illustrations are not "easy", "automatic", or "universal" sources of information is to be found in the literature on cross-cultural variation in the interpretation of pictorial information. For example, Hudson (1960) has noted cultural variation in the interpretation of perspective cues between Western cultures and various African sub-cultures. Similarly, Deregowski (1972) has reported that Zulu's are not susceptible to the Müller-Lyer illusion, nor are they susceptible to other illusions that depend on right angle cues. He explains this finding by observing that Zulu's have little exposure to right angles in their environment and thus are not misled by the arrow cues. Although such studies can be criticized on the basis that the particular differences observed may be confounded with language and other cultural differences, such observations are at least suggestive that how illustrations are interpreted is learned through experience.

Training studies. Peeck (1993) takes the logical position that one way to overcome the possibility that readers tend to underuse pictorial information is to explicitly train readers to do so. Peeck reviews the research on adding instructions to illustrated text. He reports that, taken as a whole, this work demonstrates equivocal support for the value of adding instructions to illustrated text. However, he goes on to suggest that the mixed results may be due to the lack of specificity of instructions given. Peeck maintains that large learning gains could be made by encouraging readers to *actively* manipulate illustration information.

Reinking, Hayes, and McEneaney (1988) examined the effects of explicitly cued graphic aids on good and poor readers. The authors found that explicit cueing increased both groups of readers attention to graphic aids and thus their ability to recall information displayed through illustrations that were redundant with text information.

Moore (1993) carried out a study in which he attempted to train high school students to process illustrated text effectively. The training program is an extension of Palincsar and Brown's reciprocal teaching model (1984) and is aimed at fostering students metacognitive processing of spatial aids. The program is called SLIC and focuses on teaching the following skills: how to Summarize the aid, how to Link the aid to the text, how to Imagine the aid, and how to Check for understanding. Moore's results demonstrated that students who were trained in how to process a map with this model remembered more information than untrained students.

Thorndyke and Stasz (1980) performed a training study in which subjects used one of three sets of procedures for learning map information. One group received training in "effective" procedures that had been related to successful task performance in a prior study. A second group was trained on a set of "neutral" procedures that were unrelated to learning success. Finally, a third group received no specific training and were instructed to use their own techniques. Comparisons of performance before and after training revealed that the group trained in the "effective" procedures demonstrated larger gains and outperformed the remaining groups in their ability to reproduce map information. Specifically the learning gains were obtained for the recall of spatial attributes, however, no group differences were found for the recall of verbal attributes. Interestingly, the success of the "effective" procedures training was limited by the subject's spatial ability such that only those subjects who demonstrated high spatial ability benefited from the training.

Summary of literature on variation in processing illustrated text. The literature on individual and group differences in reading illustrated text demonstrates that readers vary in their pattern of using illustrations while reading (e.g., Beveridge and Griffiths, 1987; Kuntz, Drewniak, & Schott, 1992; O'Brien & Albrecht , 1992; Reid & Beveridge, 1990). These difference have been related to levels of reader ability and reader prior knowledge and experience. In particular, it has been suggested that low ability and low prior knowledge readers make inefficient and ineffective use of illustrations and are therefore more prone to experience the presence of text illustrations as inhibitory rather than facilitory. However, little work has been conducted which attempts to explicitly link such apparent processing differences to on-line measures of the content of comprehension. Thus, little is known about how the reader's pattern of illustration use specifically affects comprehension processing.

Although, researchers in this area often suggest that low ability and low prior knowledge readers are likely to benefit from explicit instruction in how to use illustrations to support learning and comprehension of text information, this issue has just begun to be explicitly investigated (e.g., Moore, 1993; Reinking, Hayes, & McEneaney, 1988; Thorndyke & Stasz, 1980). However, with the exception of the study by Thorndyke and Stasz, researchers have yet to investigate how readers are processing illustrated text. Thus, it would be interesting to more clearly assess the effects of natural individual differences of on-line processing patterns on comprehension performance.

Reciprocal Facilitation Effects of Illustrations and Text on Comprehension

While most research has focused on the question of whether adding illustrations to text can facilitate the comprehension and learning of text content, there is also evidence that text can aid in the processing and interpretation of illustration information. Several studies demonstrating such effects are reviewed below.

Bower, Karlin, and Dueck (1975) conducted a similar study to that of Bransford and Johnson (1972). However, rather than studying the disambiguating effects of illustrations on the processing of text, Bower et al. were interested in assessing possible disambiguating effects of text on the processing of obscure illustrations. They report two experiments in which they investigated the effects of text sentences on peoples memory for "nonsensical" illustrations (i.e., droodles). In the first study, memory for ambiguous pictures was compared between conditions in which the illustrations were and were not accompanied by textual descriptions. Free recall was found to be significantly superior in the illustration with text condition. In addition, subjects who received the illustration with text were more likely to confuse similar distracter illustrations with the original on a recognition memory task. In the second experiment, subjects studied pairs of ambiguous illustrations with and without an accompanying phrase which identified and interrelated the illustrations. Subjects who were presented with both illustrations and text showed superior associative memory performance. The authors concluded that, similar to the effects of illustrations on the processing of ambiguous text, memory for illustrations can be improved by evoking appropriate schemata through textual descriptions.

Jörg and Hörmann (1978) studied the influence of the level of specificity of verbal labeling on the depth of subsequent picture processing. These researchers assigned readers to one of three experimental groups. One group read a sentence which described two of four depicted objects using specific labels (e.g., tulip) prior to viewing the accompanying drawing. A second group read a sentence which also described two of the four depicted objects, but used general labels (e.g., flower), prior to viewing the same drawing. A third control group was simply exposed to the drawings without any accompanying text. Results demonstrated that the generality of the verbal descriptions affected subjects' recognition memory performance for *both* labeled (i.e., textually described) and unlabeled objects. The authors concluded that the specificity of terms used to label objects affects the depth at which readers process subsequently presented pictorial information. Thus, Jörg and Hörmann's data provide clear support the notion that text features (i.e., the level of descriptive specificity) can affect the manner in which picture information is subsequently processed.

Guri-Rozenblit (1988) investigated the differential processing of verbally explained and unexplained abstract diagrams embedded within social science texts among adult readers. Participants included 416 university students who were assigned to read one of two texts (one text was about the juvenile court system and the other was about marketing communication) of roughly 4,000 words. For each text domain, male and female subjects were assigned to one of four conditions: a text only control, a text plus unexplained diagram, a text plus verbally explained diagram, or an elaborated text with no diagram. Comprehension and learning were assessed using multiple choice items, and open ended questions which involved both verbal and graphic tasks. In addition, standardized verbal and visual aptitude tests were administered to be used as covariates. The main results of the study indicated that a verbally explained diagram is more effective than an unexplained one in supporting readers' ability to understand complex explanations; a diagram, in general, is more effective than a verbal explanation at representing sequential and hierarchical relations; diagrams have a significant facilitory effect on both 'active' recollection and 'passive' retention; the mode of the text's design and presentation is more influential than the reader's initial verbal and visual aptitudes, and there is no gender difference in processing of verbal and diagrammatic representations. Thus, Guri-Rozenblit's results provide further support for the notion that verbal descriptions (i.e., text) can facilitate the comprehension of complex diagrams.

Greenspan and Segal (1984) also report data that supports the idea that text features can affect the processing of illustrated information. Using a sentence-picture verification priming task these researchers found that sentences can strongly influence the learner's interpretation of visible nonlinguistic events (i.e., illustrations) and that this representation is sensitive to the distinction between textually presupposed (i.e., topic) and asserted (i.e., comment) information. According to the authors, such results suggest that the topic of a discourse is not merely a linguistic entity but rather can be thought of as a an entity that is established in a nonlinguistic (e.g., an illustration) context through discourse.

Researchers who adopt the text appropriate processing view of understanding illustrated text have also concerned themselves with the study of how the characteristics of text and illustrations may interact to affect overall processing. The text appropriate processing hypothesis is based on the assumption that different text structures induce different patterns of cognitive processing. According to this position, narratives which describe a series of temporally and causally related events tend to evoke relational processing in readers, while expositions which typically describe objects and their properties, tend to evoke individual item processing. When extended to the case of illustrated text, the hypothesis states that the effects of illustrations on text processing will depend upon the relationship between the type of information depicted by illustrations and the type of text in which illustrated information is embedded. Thus, this position focuses on the relationship between illustration and text properties.

For example, Waddill, McDaniel, and Einstein (1988) have presented evidence that the effects of illustration on memory for text varies as a function of both text type (narrative versus expository) and as a function of the type of information depicted by illustrations (details versus relations).

Rusted and Hodgson (1985) have also demonstrated that the nature of the facilitory effects of adding illustrations to text depends to some degree on the text type. Rusted and Hodgson assigned forty nine-year-old children to either a factual (expository) or a fictitious (narrative) passage condition which was either accompanied by an illustration or not. Immediate free recalls were collected as an index of learning. Results revealed an overall picture facilitation effect, however, the nature of facilitation depended on text type such that increases in recall were limited to information that was directly illustrated in the narrative text, but were observed for both illustrated and unillustrated content in the expository text condition. Furthermore, recall for unillustrated narrative text content was found to be lower for subjects in the illustrated versus unillustrated narrative condition. The authors interpret these results as support for the notion that text illustration in narratives serve to emphasize illustrated information and tend to draw attention away from unillustrated information. Thus, the presence of illustrations in narratives may interfere with the reader's construction of a hierarchical representation of the story. However, for expository texts, it appears that no such interference occurs. On the contrary, the addition of illustrations to expository texts seem to enable more effective processing of all text content.

Although the text appropriate processing argument is intriguing, it's conclusions are in direct contrast to existing work demonstrating that adding relevant illustrations to narrative texts does indeed facilitate memory for text content, including unillustrated text content (e.g., Bluth, 1972). Furthermore, several studies have failed to find significant increases in readers' memory of unillustrated information when they are provided with illustrated expository text (e.g., Peeck, 1980). In particular, this argument appears to be somewhat simplistic, and may have more to do with the semantic complexity of the information in the two texts as opposed to the information type. That is, no explicit attempt has been made to investigate whether such effects recur when texts are balanced for informational complexity. However, the position is important for pointing out that the effects of illustrations on text processing are not necessarily general.

Summary of reciprocal facilitation effects. While some researchers have focused on demonstrating that illustrations can facilitate text processing, others have focused on the facilitory effects of text on illustration processing. The point to be made here is that when readers are presented with a combination of related information sources, neither medium is interpreted by the reader in isolation. Clearly, illustrations can affect text processing and vice versa. This observation emphasizes the need to study the reciprocal effects that each modality of information may exert on the other(s). Furthermore, once one adopts such a position, several pertinent questions suggest themselves. For example: How does the reader's use and understanding of one source of information affect the manner in which they use and understand information presented in an alternative medium? How does the reader resolve differences in terms of how information is organized across media? How should text-illustration combinations be designed and sequenced to stimulate integrative processing across media? If we are to come to an adequate understanding of how illustrated texts are processed, we must explicitly address such questions.

The discussion now turns to consideration of the issue of knowledge integration. In particular, the concepts of coherence, cohesion, and knowledge integration as they have been studied in the area of text comprehension are reviewed.

Coherence & Cohesion

It is widely recognized that discourse comprehension requires the ability to construct a coherent integrated representation of that discourse. Skilled readers are able to quickly and efficiently integrate information from various sources as they build a representation of the text (Anderson, 1993). The question of how coherence links are generated, although related to surface structure features, transcends the linguistic aspects of a discourse and has been characterized as being cognitive in nature. Text coherence refers to the extent to which concepts are sequenced and presented in such a way as to make the relationship between ideas transparent to the reader. The perception of coherence is achieved through the use of a variety of features which connect concepts, some of which include: the prior general knowledge of the reader and involves the use of inference, deduction, and presupposition; the use of vocabulary; the use of punctuation and layout (e.g., panels, headings, bullets, etc.); and may also be signaled in spoken discourse through prosodic features (i.e., variation in pitch, loudness, speed, rhythm, and pauses). The local and global coherence of a discourse is expressed and marked by surface properties such as clause organization, clause ordering, sentence ordering, connectives, pronouns, adverbs, verb tenses, lexical identity, paraphrases, and definite articles (van Dijk, 1985). Research on text coherence generally assesses the degree to which there is repeated reference to the same set of entities (e.g., Garnham, Oakhill, & Johnson-Laird, 1982; Kintsch & van Dijk, 1978).

While the property of cohesion has also been described as being semantic in nature, it is distinguished from the concept of coherence in that it refers specifically to relations of meaning that exist *within* a text. Cohesion occurs when the interpretation of some discourse element is dependent on that of another. Thus, cohesion is viewed as one of the linguistic aspects which contribute to the reader's ability to experience a text as coherent (Moe, 1979). Halliday and Hasan (1976), provide a qualitative description in which text cohesion depends on a set of connections between words and concepts in different sentences. Their system of analysis distinguishes cohesive relationships that are signaled through reference, lexical cohesion, conjunction, substitution, and ellipsis.

Studies have demonstrated a positive relationship between a texts cohesion and/or coherence and its comprehensibility (e.g., Anderson & Armbruster, 1984; van Dijk & Kintsch, 1983; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Marshall & Glock, 1978-79; McKeown, Beck, Sinatra, & Loxterman, 1992). Specifically, the effects of text cohesion and coherence appear to affect the readers ability to relate information across sentences and paragraphs. For example, Irwin (1980) studied the relationship between the number of cohesive ties (defined according to Halliday and Hasan's system) and free and prompted recall. She compared college student's recall performance for two texts about gibbons. One text had about twice as many cohesive ties as the other. Irwin found that although there were no differences between groups in their ability to recall micro-level propositions, there was a significant difference between the groups on their ability to recall macro-level propositions. Thus, it appears as though cohesive ties facilitate readers' ability to generate a global connected representation of the text.

It is also true, however, that authors do not typically include a complete description of all relevant information and their relationships in their prose. In such cases, the reader is required to form inferences to fill in gaps and to connect relevant information that is necessary to form a coherent representation of the text. Evidence that readers' routinely generate such integrative inferences as part of the comprehension process are plentiful. For example, Kintsch (1974) presented readers with brief passages that either required inferences (implicit version) or did not (explicit version) and tested their memory on a verification task either immediately or following a 15 minute delay. The results demonstrated that readers were able to distinguish implicit from explicit information when tested immediately, but tended to confuse sentences they actually read (explicitly presented information) with implicit versions following a brief delay. Kintsch concluded that memory for text involves at least two distinct representations: a short lived verbatim or surface structure representation, and a more enduring propositional representation which includes an integration of both explicitly presented information and reader generated inferences.

It is interesting to note that similar effects have been demonstrated for the comprehension of picture stories. For example, Baggett (1975) carried out a similar experiment in which subjects were provided with a series of pictures depicting common events. Subjects were given a verification task on either propositions that were directly represented in the pictures or were inferable from the pictures. Again, readers were able to distinguish explicit from implicit information immediately after stimulus presentation and following brief delays, but tended to confuse such information following a 72 hour delay. These results are similar to Kintsch's (1974), with the exception that this effect required a 72 hour delay instead of a 15 minute delay. Such results demonstrate that the human cognitive system is biased towards attempts to construct an integrated coherent representation of information expressed through either text or illustrations.

It has also been suggested that the comprehension benefits of a highly coherent text may be different for readers who vary in terms of their prior knowledge of the text content domain. For example, McNamara, Kintsch, Butler Songer, and Kintsch (1996) studied the effects of varying the local and global coherence (defined in terms of argument overlap) of biology texts on junior high school students comprehension at both the textbase and more abstracted situation model levels of understanding. Their results indicate that coherence contributes fundamentally to low prior knowledge readers' ability to form an adequate textbase representation, but did not affect high prior knowledge readers' ability to do so. However, highly coherent text appeared to interfere with high prior knowledge readers tendency to construct a situation model of the text. The authors argue that minimally coherent texts may induce high prior knowledge readers to more *actively* process the text which in turn facilitates their ability to construct a high level understanding of the material.

Summary of coherence and cohesion. A texts level of coherence and cohesion affects the ability of readers to build a connected representation of the text content. Furthermore, these properties have been shown to affect different levels of discourse representation (i.e., micro versus macro structure, and textbase versus situation model) and are different for reader's who vary in their prior knowledge (i.e., a high degree of coherence may benefit the ability of low prior knowledge readers to construct an adequate textbase representation, but may interfere with the ability of high prior knowledge readers to form a situation model). Such results support the contention that coherence is not simply a property of the text, but also involves active constructive processes on the part of the reader.

In addition to the effects of text cohesion and coherence on comprehension, there is evidence that readers are also sensitive to cohesion and coherence relations expressed through illustrations. Such results support the idea that readers attempt to integrate related information into a coherent representation in each medium. Given these observations the question of whether readers are sensitive to coherence and cohesion relation across media as well as how such relationships may be signaled to the reader arises. We turn now to the topic of knowledge integration proper as it has been studied in the areas text processing and illustrated text processing.

Knowledge Integration

Knowledge integration and text processing. As stated above, one of the tasks a reader faces when he/she reads a text for the purposes of learning and understanding is to integrate information into a coherent representation. Integrated representations are of value in that they enable the reader to simultaneously consider related information and can thereby facilitate high order cognitive processes such as summarizing, inferencing, reasoning, and decision making (Walker & Meyer, 1980a). In the case of understanding text, two kinds of knowledge integration have been distinguished; integration of information expressed within a text, and integration of text-derived information with the reader's prior knowledge (Frederiksen & Breuleux, 1990; Potts, 1977).

With regard to the first kind of knowledge integration, various surface features of the text can affect the likelihood that conceptually related text information will be integrated by the reader. For example, integration is more likely to occur when there is a high degree of correspondence in the wording of related information, and when relational links between concepts are expressed closely together in the text than when they are expressed distally (Hayes-Roth & Thorndyke, 1979). Thus, to some degree, the manner by which coherence of the text is explicitly marked can affect the readers ability to integrate information into a connected representation.

Other text structure features may also influence the likelihood that integration of text information will occur. For example, Walker and Meyer (1980b) investigated whether integration of text information depends on the height of information in the text structure and whether it is possible to differentiate between integrative processes that take place during information acquisition (structural integration) from that occurring at retrieval. Using an inference verification task, Walker and Meyer found that, regardless of whether subjects were instructed to learn or to simply read the material, information that occurred high in the text hierarchy was more likely to be integrated than low level information. In addition, verification times were significantly faster for information which occurred consecutively (i.e., evidence for structural integration) than information which occurred separately. These results suggest that information integration is affected by both the conceptual height of information in the text as well as its relative location.

Staging is yet another text variable which has been shown to influence the integration and memory of text information (e.g., Britton, Glynn, Meyer, & Penland, 1982; Clements, 1979; Kieras, 1981; Marshall & Glock, 1978). Grimes (1975) defines staging as the hierarchical organization of text propositions as being superordinate or subordinate to one another as expressed by the text base. As such, staging is a dimension of text structure which reflects the prominence given to various segments of the text within a discourse.

In addition to integrative operations that occur within a text, research has demonstrated that there are distinct processes associated with integrating text-derived information with the reader's existing prior knowledge. Such integration processes include a variety of retrieval, inferential and reasoning processes that operate in using prior knowledge to understand or interpret new information. For example, Mannes (1994) describes a "reinstatement-and-integration" strategy which focuses on the integration of knowledge derived from separate text sources (e.g., a text and an outline). According to this model, previously processed and stored information is activated and reinstated in the reader's short term memory buffer whenever a coherence break is detected or a previously presented topic is encountered. When a reinstatement occurs, so does the potential for knowledge integration since the contents of the short term memory buffer will contain both the previously stored information and new text-derived information.

Kubes (1988) studied the ability of chemistry experts and novices to integrate relevant prior knowledge with new derived information from a series of texts about photosynthesis. She also examined the effects of task cues on the accessibility and use of prior knowledge. Kubes found that both prior knowledge and literal comprehension contributed significantly to predicting the likelihood that knowledge integration would take place. Such research demonstrates that knowledge integration is not necessarily automatically evoked, even when appropriate. Rather, it appears that integrative processes depend on both aspects of the task environment and readers' level of domain relevant prior knowledge.

Potts, St. John, and Kirson (1989) demonstrated that the degree to which readers integrate new text derived information with their prior knowledge differs depending on whether they are led to believe the new information is real or artificial. Thus, in addition to the finding that knowledge integration depends on the reader's prior knowledge, there is also evidence to suggest that integration depends on the reader's perception of the task materials.

Past research has demonstrated that the task of self-explanation while studying worked examples can facilitate the learner's ability to integrate newly acquired information with prior knowledge in acquiring problem-solving skills (e.g., Chi, Bassok, Lewis, Reimann, & Glasser, 1989; Chi & VanLehn, 1991; Ferguson-Hessler & de Jong, 1990). More recently, Chi, de Leeuw, Chiu, and LaVancher (1994) have extended the study of this phenomenon from the effects of spontaneous self-explanations on the ability to learn procedural information from worked examples, to the study of the effects of explicitly promoted self-explanation on the ability to learn declarative knowledge from expository text. In this study, 14 eighth-grade untrained students were asked to self-explain each line of a text describing the functioning of the human circulatory system as they read. Ten students in a control group simply read the text twice without any instructions to provide self-explanations. Chi et al. found that students who were explicitly instructed to self-explain demonstrated better learning gains on a posttest than the non prompted group. Furthermore, of the self-explanation group, students who demonstrated a large number of explanations (designated as "high" explainers) outperformed the remaining students when asked to answer very complex questions, and were better able to induce the function of components from their understanding of the system. Mental model analysis of self-explanation protocols also revealed that each of the high explainers achieved the correct mental model of the circulatory system, while many of the remaining students did not. The authors concluded that the act of self explanation facilitates the ability of readers to integrate text-derived information with their prior knowledge and the construction of appropriate mental model representations.

Knowledge integration and processing of illustrations and text. Knowledge integration effects are not limited to prose. In addition to Baggett's (1975) study mentioned above, Franks and Bransford (1971) provided an early demonstration that peoples memory for visual patterns tends to be abstracted and integrated as well. These researchers presented subjects with a set of geometric figures that represented varying degrees of transformations of a visual prototype. During an acquisition phase of the experiment, participants were presented with a subset of transformed examples of the prototype target concept. After acquisition, participants were tested on their ability to recognize new examples of the concepts acquired. The results indicated that when people were asked to identify new examples of the concepts, they chose those that were closest to the prototype, even though the prototype itself had never been presented. These data were interpreted as evidence that people abstract and integrate visual information.

In the case of trying to comprehend illustrated text, the complexity of the reader's task is compounded. In such situations, the reader is additionally faced with the task of integrating information expressed by one or more illustration(s), and integrating that information with text-derived information. Baggett (1989) has also pointed out that in the case of dual media presentation, there are at least two types of cohesion to be considered - within media and between media. However, we currently do not know exactly how readers select and coordinate information across media. For example, it is not at all clear what information should be illustrated to encourage integrative processing. Nor is it clear how information should be sequenced (both within and across media) to promote integrative processing.

Yee, Hunt, and Pellegrino (1991) performed several experiments demonstrating separate individual difference and task effects on the ability to coordinate perceptual and verbal information. The authors claim that their results demonstrate that the ability to integrate information across media is a distinct ability from that of integrating information within each medium.

Rather than making the process of integration more difficult, Winn (1987) has argued that illustrations have the potential to facilitate integration. Specifically, he has claimed that illustrations may decrease working memory demands and thereby enable the learner's limited cognitive resources to be devoted at higher order operations such as developing a coherent semantic macrostructure. However, it is not clear from Winn's position whether he is talking specifically about the integration of textual material or whether he is referring to cross-modal integration.

Chandler and Sweller (1991) report a series of experiments in which they investigated the effects of various formats for presenting text and diagrams on learning. The authors argue that, for situations in which both textual and diagrammatic information are necessary to learning, the manner in which information is presented will affect the cognitive load of the learner such that cross-modal information presented distally will exert undue cognitive load and interfere with learning. On the other hand, information presented in an integrated format (e.g., worked examples) will appropriately focus attention and reduce cognitive load, thereby enabling learners to devote their limited cognitive resources to the task of learning. The reader may note some similarities here between this position and Mayer's (1994; 1997) contiguity principle. However, Mayer's position is distinguished by the hypothesis that presenting cross-modal information distally will specifically denigrate low prior knowledge learners ability to form cross-modal referential relations and thereby their ability to construct a runnable mental model. On the other hand, Sweller and his colleagues make no specific claims about what kinds of representation or processes are specifically or differentially affected by a heavy cognitive load.

Hegarty and Just (1993) report two experiments designed to assess readers' comprehension and eye fixations as they read increasingly complex descriptions of pulley systems. Hegarty and Just hypothesized that people attempt to integrate information from text and diagrams rather than form two separate representations. In order to test this hypothesis, their first study was specifically designed to assess the individual and conjoint effects of texts and diagrams on readers' ability to construct a mental model of a pulley system. Subjects of either high or low mechanical ability were presented with either a text alone, a set of diagrams alone, or a combination of text and diagrams describing three increasingly complex pulley systems. The texts described both configuration and kinematic information, while the diagrams conveyed only configuration information. Comprehension of both configuration and kinematic information was assessed using a set of post-reading questions that were designed to assess the accuracy of subjects' mental model representations of this information.

Hegarty and Just found that, overall there was a significant main effect of information type on mental model comprehension such that configuration information (available from both the texts and illustrations)was better comprehended than kinematic information (available in the illustrations only). In addition, they found a significant interaction between the medium of presentation and information type such that the comprehension of kinematic information was facilitated for subjects who read the illustrated text combination. With regard to the comprehension of configuration information (available from both the text and the diagrams), subjects who received the text and diagram combination outperformed the diagram only condition, but did not differ from the text only condition. This result supports the conclusion that subjects can construct a representation of configuration information from a text description alone. Taken as a whole, this pattern of results also supports the notion that the effects of text and illustration are reciprocal. In particular, the finding that the text and diagram combination was particularly advantageous at promoting the comprehension of kinematic information (presented in the text only) supports the notion that illustrations affect the processing of text information, and the finding that the text and diagram combination comprehended configuration information (available in both text and illustrations) better than the diagram only group suggests that text can facilitate the processing of illustration information. Taken as a whole, this pattern of results suggests that readers do indeed integrate information across media.

While it was expected that high ability readers would outperform low ability readers in all conditions, statistical differences were only found for the text and diagram combination condition. With regard to the comprehension of kinematic information (available from the text only), the text and diagram combination proved to be more facilitative than either a text or diagram alone. No differences were found between the text only and diagram only conditions. Furthermore, high ability subjects comprehended kinematic information better than low ability subjects for the text and diagram combination, and the text only conditions, but did not differ for the diagram only condition. Based on these results, the authors concluded that a combination of text and diagrams facilitates the reader's ability to construct an integrated mental model, particularly with regard to the comprehension of kinematic information.

In their second study, Hegarty and Just proposed that some aspects of mental model construction could be inferred from readers' pattern of inspection of text and related illustrations. Using the same stimulus materials as in study 1, they measured the eye fixations of 5 high mechanical ability and 4 low mechanical ability subjects. In addition, they measured subject's mental model comprehension with a subset of the open-ended comprehension questions employed in study 1. Overall, comprehension of configuration information (presented in both text and diagrams) was found to be better than comprehension of kinematic information (presented in the text only). While no overall comprehension differences were found with regard to mechanical ability were observed, the manner in which high and low prior knowledge readers accessed the materials did differ. Low mechanical ability readers required longer study times and made more frequent text regressions and diagram inspections. Overall, integration across media, which was operationalized as points where subjects interrupted their reading of the text in order to inspect the diagram, was found to increase with the complexity of the pulley system described, and to occur primarily at the end of a clause or sentence. Hegarty and Just concluded that: (a) readers need to process both media to form a complete mental model, (b) readers integrate information from text and diagrams incrementally from a local representation of several components to a global representation of the entire system, and (c) although low mechanical ability readers are able to form an integrated mental model representation from illustrated text, they require longer study times and more frequent text regressions and diagram inspections than high mechanical ability readers.

Summary. Comprehension of either text or illustrations involves the ability to construct a coherent representation of conceptual information. The construction of such a representation involves the use of both abstractive and integrative processes. In addition, the readers ability to integrate information presented in either medium appears to be affected by similar characteristics (e.g., prior knowledge, argument overlap, and distance between related information).

The problem of understanding illustrated text, or any multimedia presentation, is compounded by the requirement that readers must not only integrate information within each medium, but must also be able to integrate information across media. Exactly how readers are able to accomplish this is not well understood. Specifically, there is a need to study how readers make use of text and accompanying illustrations as well as what readers are understanding as they read.

Based on literature reviewed above, it is apparent that the problem of understanding and learning from illustrated text is complex and involves integrated considerations about illustration, text, reader, and task characteristics. Furthermore, adequate study of the problem requires the ability to accurately describe the information content of illustrations and texts. We turn now to consideration of cognitive models of discourse processing which provide a well defined integrated set of assumptions about how such variables influence comprehension processing as well as a host of methodologies for assessing such effects . Indeed, Kintsch (1998) has made a convincing argument for using comprehension as a general unifying paradigm for studying cognitive phenomena.

Cognitive Models of Discourse Processing

Discourse comprehension is currently viewed as a highly inferential and interactive cognitive activity in which the reader constructs multiple representations simultaneously. It is interactive in two senses. First, comprehension involves the use of both top-down or knowledge-driven processes and bottom-up or text-driven processes (e.g., Carpenter, Miyake, & Just, 1995; Frederiksen, Bracewell, Breuleux, & Renaud, 1989; Just & Carpenter, 1987; van Dijk & Kintsch, 1983; Perfetti, 1985; Rumelhart, 1977). That is, comprehension is seen a complex cognitive process in which readers actively construct meaning on the basis of information in memory and new information derived from the text. Secondly, it is computationally interactive in that processing at one level of representation can affect processing at other levels (e.g., Carpenter, Miyake, & Just, 1995; Frederiksen, Bracewell, Breuleux, & Renaud, 1989).

The cognitive perspective on discourse processing places great emphasis on addressing the complexity of comprehension by attempting to describe and understand the cognitive representations and component processes that readers use in trying to understand a text. For example, van Dijk and Kintsch (1983) distinguish between three distinct levels of representation. The *surface structure* captures the exact wording of a text and is the most superficial and short lived level of representation. Construction of a surface level representation involves the use of highly automatic lexical and syntactic processes. The *textbase* represents the semantic content of the text itself as a network of connected propositions. Finally, the *situation model* represents the situation or state of affairs described by the text, but is distinct from the text itself. This level of representation results from processes that integrate text derived information with the readers prior knowledge. Experimental research has consistently supported the psychological validity of these distinct levels of representation (e.g., Fletcher & Chrysler, 1990; Schmalhofer & Glavanov, 1986; van Dijk & Kintsch, 1983; but see Fletcher, 1994 for a comprehensive review).

Similarly, Frederiksen and his colleagues (e.g., Frederiksen, Bracewell, Breuleux, & Renaud, 1989; Frederiksen, 1986; Frederiksen & Donin, 1991), present a detailed model in which text understanding is viewed as a stratified modular process in which several types of representations are generated simultaneously by the reader. These researchers have described discourse comprehension as involving the following levels of representation.

At the linguistic level, morphemes and words are represented and provide the input for the generation of syntactic parse tree structures, as well as the representation of thematic and referential relations which serve to link syntactic structures, (e.g., the resolution of anaphoric reference).

At the local conceptual level, *propositions* are generated from syntactic and morphological input. Propositions represent elementary truth-valued semantic units that can be directly expressed in language and which serve as a basis for the acquisition of conceptual knowledge structures. Current propositional models (e.g., Frederiksen, 1975; 1981) are capable of representing a variety of semantic relations and structures including: (a) the description of resultive actions (i.e., events and associated case and identifying relations), (b) the identification, determination, and quantification of objects (i.e., stative relations), (c) the description of stative and resultive processes (i.e., systems) (d) properties of abstract concepts representing propositions (i.e., propositional relations), (e) relations linking concepts or propositions into identity sets (i.e., identity relations), (f) the identification of operations defined on operands that return values (i.e., functional relations), (g) relations that describe how one proposition depends on another (i.e., binary dependency relations), and (h) relations describing the alternative or exclusivity of propositions (i.e., conjoint dependency relations).

At the global conceptual level, *conceptual graphs* in the form of semantic networks represent interconnected conceptual information reflected in propositions. In addition, broader types of conceptual networks, such as procedures, narratives, descriptions, etc., may be generated to represent specialized types of semantic information. Such structures fall under the labels of *scripts* (Schank & Abelson, 1977), *schemata* (Rumelhart, 1980), *situation models* (van Dijk & Kintsch, 1983), *mental models* (Johnson-Laird, 1983) and *frames* (Frederiksen, 1986). It is these structures which may be used to integrate text-derived information and the reader's prior knowledge structures.

Thus, the comprehension of text involves the construction of propositions from the text natural language input, as well as the integration of individual propositions into connected network structures which may themselves be embedded within specialized knowledge structures (Frederiksen, Bracewell, Breuleux, & Renaud, 1989). Discourse comprehension, then, is viewed as a process where textual and linguistic structures are used in conjunction with the reader's prior knowledge to *construct* an abstract conceptual representation of the text (van Dijk & Kintsch, 1983; Frederiksen, 1986).

Memory for text vs. learning from text. Kintsch and his colleagues have made an effort to distinguish between readers' ability to *remember* text, which involves the construction of a textbase representation, and their ability to *learn* from text, which involves the construction of a situation model (e.g., Kintsch, 1994; Mannes & Kintsch, 1987; Perrig & Kintsch, 1985). Kintsch (1994) has emphasized that this distinction between memory and learning has strong implications for the design of texts. That is, text structures which facilitate readers' memory for text do not necessarily facilitate their ability to learn from text. Thus, he warns that we need to be concerned with specifying exactly what our instructional goals are when designing texts. Specifically, Kintsch has argued that we need to be concerned with answering the following questions: "How is learning affected by the content of texts and by the form of texts?". Clearly, these questions should also be of concern when applied to the situation of designing illustrated texts.

Methods and measures based on current perspectives. Considerable progress has been made over the last 20 years in understanding the moment-by-moment dynamics of discourse processing. These advances are largely due to the development and use of multiple measures and methods of assessing comprehension that are grounded in cognitive models of discourse processing.

Traditionally, discourse research was limited to employing post-input measures of comprehension including, free recall, various cued recall, and recognition measures. Of these post-input measures, oral recall has been generally preferred because it provides evidence of what information was processed and could be retrieved by the reader. Oral recalls are considered to be relatively unbiased estimates of comprehension in the sense that they do not provide readers with memory cues. Furthermore, recall measures may provide information about what inferences the reader made as well as how the reader organized their representation of text information. However, because a recall task is temporally removed from the reading task itself and necessarily involves both comprehension and retrieval processes, there is a possibility that readers may not include all information that they comprehended and are able to remember (i.e., retrieval during recall may not be exhaustive or may be distorted). Similarly, recall data do not necessarily provide evidence of all processes that may have occurred during reading. That is, recall is a product rather than a process measure.

To get around this issue, researchers began to employ various on-line measures, including reading times, frequency and duration of eye-fixations, and concurrent verbal reports to gather data relevant to how readers employ comprehension processes in real time. For example, reading time is assumed to index how much processing is going on (i.e., where processing load is heavy). Whereas, concurrent protocols or an on-line interpretation task provide information about what processes are operating and allow the researcher to obtain evidence about when a given process is used in reading. However, on-line measures alone may not be sensitive to overall effects of comprehension processes. For example, the overall memory structure of the representation may not be indexed by such measures. Thus, using indices of both on-line and post-input allows researchers to obtain concurrent evidence that a given process has occurred, its temporal locus, and its effect on overall comprehension and retrieval processes.

The point is to be made here is that it is now widely recognized that discourse processing ought be studied using multiple measure of comprehension, including both on-line and post-input measures.

Discourse from a situated cognition perspective. Recently, the cognitive analysis of discourse has been broadened to consider the ways in which characteristics of the surrounding context contribute to discourse processing and learning. Theorists within this framework have criticized the direct (i.e., unaltered) extension of the "learning from text" paradigm as an appropriate model to the study of all kinds of learning. It is argued that the "learning from text" framework looses sight of the fact that discourse is not an isolated phenomenon, but rather is intimately and functionally tied to some situational context. According to such a perspective, face to face communication occurring within in a particular context is viewed as the most natural human learning situation. Alternative learning situations (e.g., learning from text, learning from illustrations, computer-based instruction, etc.) are viewed and understood as variations on this basic situation in which various devices and conventions have been adopted to compensate for such modifications. Indeed, one major goal within this framework is to gain a better understanding of how learning in these alternative situations takes place and how it may differ from one kind of situation to another.

Such a position has led to a broadened view of comprehension processing that extends beyond language. For example, Greeno (1989) has described a perspective that emphasizes the role of mental models for integrating multiple kinds of representations within a discipline. In a similar vane, van Dijk (1985) positions semantics within the broader scope of semiotics. He argues that not only is there a semantics of natural language utterances and acts, but also of non verbal or paraverbal behavior, including gestures, pictures, films, logical systems, sign languages, and social interaction in general. By the same token, however, many theorists still place discourse at the center of their models.

For example, Frederiksen and Donin (1996) have described a model of situated discourse that they have developed to study the role that discourse plays in learning in various complex interactive situations (e.g., one-on-one tutoring , small groups interactions, and large science classrooms). According to these researchers the term "situated discourse" refers to any discourse that is bound to the context in which it occurs. Thus, "learning through situated discourse" is to be distinguished from the traditional "learning from text" perspective in that the interpretation of the discourse is viewed as constrained by characteristics of the surrounding context that extend beyond those by traditional author - reader - text models. Furthermore, it is assumed that the study of how discourse is linked to and constrained by aspects of the surrounding situation will lead to more clearly specified general models of discourse processing and learning.

Frederiksen and Donin's model of situated discourse (see figure 1) contains the following components. At the center of the model are the constituents of discourse which are linked to one another through *endophoric* relations to form a cohesive text. The discourse in turn is linked through *exophoric* relationships to elements in the representational environment (including symbols, icons, images, expressions, etc.), elements in a shared spatio-temporal environment, and elements in a shared physical environment (e.g.., objects, states, actions, events, etc.). In addition, discourse is linked to the participants' knowledge structures (e.g., concepts, propositions, mental models, semantic networks, etc.) through *semantic* links. Discourse may also be linked to features of the social and interactional environment (e.g., conversational structure, status of participants, etc.) through *functional* relations. Finally, discourse is linked to the participants' affective states (e.g., emotions, attitudes, beliefs, etc.) through *affective* relations. All of these surrounding elements produce potential constraints on the ongoing production and interpretation of discourse.

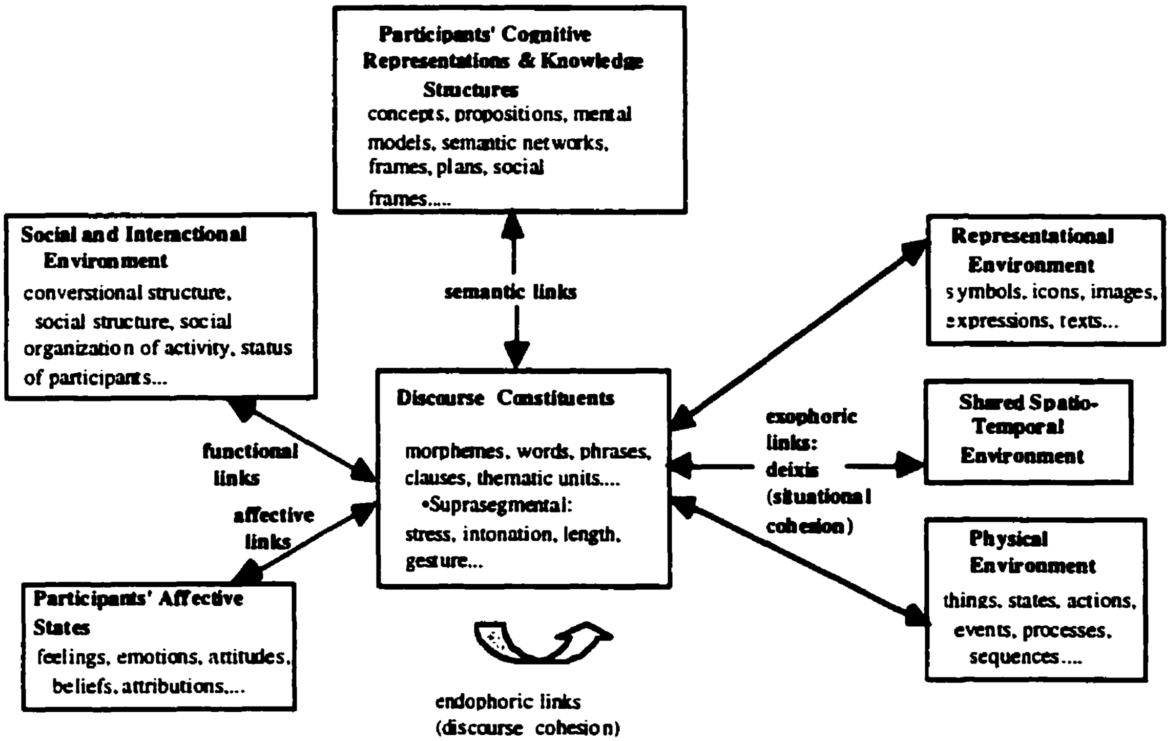


Figure 1. Model of situated discourse (reprinted with permission (Frederiksen & Donin, 1996)).

Given this extended perspective of discourse, one can view the situation of learning from illustrated text as falling somewhere between the two extreme situations of learning from text alone and learning in a face to face context. That is, illustrations may certainly be used to help contextualize the discourse of a text, and in this sense may be viewed as somewhat more "situated" than traditional texts. On the other hand, illustrated texts can probably be characterized as less situated than say learning from discourse in a classroom. Exactly how such constraints are signaled to and processed by the reader, however, is not well understood.

One of the goals of the present research is to adopt this notion of situated discourse to the case of comprehending and learning from illustrated text. By adopting such a perspective, it is fairly straightforward to see how one can extend models of cognitive discourse processing to the situation of reading illustrated text. Of particular relevance to the current problem under study is the notion that not only are there distinct processes involved in the ability to integrate text information itself and text information with prior knowledge, but there may also be distinct processes involved in the ability to integrate illustrated information as well as the ability to integrate information across media. One of the benefits offered by this model of situated discourse is that it begins to elucidate what may be meant by the terms "cross-modal cohesion" and "cross-modal integration".

Figure 2, below, depicts the main components of Frederiksen and Donin's model which are relevant to the study of illustrated text comprehension. Note that related information *within* each medium is linked through endophoric cohesive relations. Factors affecting discourse cohesion have been amply studied. Such research demonstrates a variety of ways in which text cohesion may be signaled to the reader, including: repetition, similarity in wording, topic-comment structure, staging, conceptual height of information, proximity, etc. . However, in the case of illustrations, the literature is not yet clear about how cohesion may be achieved. Although, previous research indicates that spatial proximity of related information may be an important signal, other mechanisms such as color (e.g., Dwyer, 1972; 1978) and figural similarity may also signal relatedness of depicted information. Also note in figure 2, that the cohesion *between* media is signaled through exophoric relationships. Exactly how cross-modal cohesion is achieved though has not been adequately studied. Again, however, past research can provide us with some clues. For example, there is some research which demonstrates that the more physically and/or spatially integrated the two external representations are, the more likely a reader is to relate the two modes of information (e.g., Mayer, 1997; Mayer, Dyck & Cook, 1984; Sweller et al., 1991). In addition, one might expect that comprehension integration might be more likely to occur when similar organizational structures are used in both media (e.g., similar information staging). However, such questions require further investigation before any firm conclusions can be made.

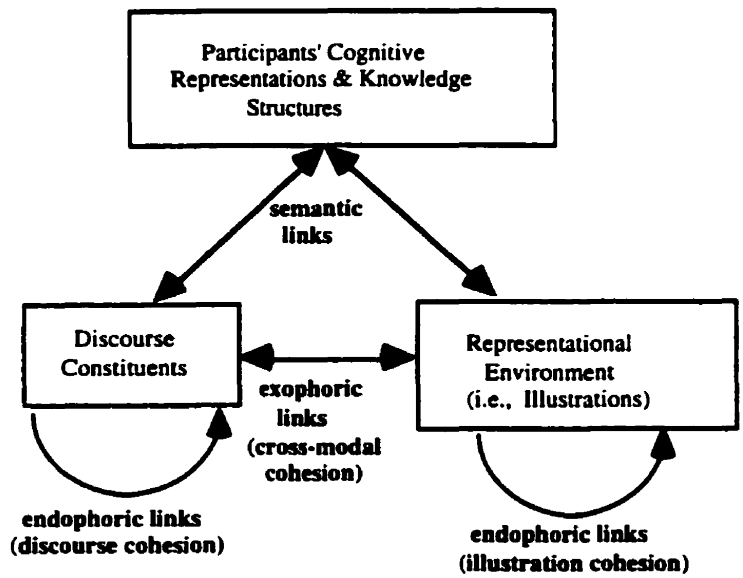


Figure 2. Frederiksen & Donin's (1996) model of situated discourse adapted to the study of illustrated text.

A strategy for investigating the cognitive processing of illustrated text. Adopting the perspectives of both cognitive models of discourse processing and situated discourse processing as described above helps to modify some assumptions about how to study the problem of understanding and learning from illustrated text that have been made in the past. First, the notion that illustrated information is "transparent" or "easy" to understand is replaced by the notion that understanding such information depends on the readers ability to construct an appropriate representation of that information, and that the reader's ability to do so may require specific knowledge. Second, the assumption that memory for text is to be equated with comprehension and learning is replaced by a more relativistic position which views comprehension as involving different levels of understanding. Third, the idea that the problem of how illustrations affect text processing can be fruitfully studied without considering how characteristics of the text may affect illustration processing is replaced by an acceptance that the effects of illustration and text on comprehension processing are likely to be reciprocal in nature. Fourth, the assumption that individual differences in terms of how readers process illustration information are not of interest is replaced by a specific interest in the variety of processing patterns that individuals may display. Fifth, the assumption that the ease of comprehension steadily increases with number of representations provided to the reader (i.e., the notion that more is necessarily better) is replaced by a realization that each source of information must not only be understood in its own right, but that related information must be integrated both across and within media. Sixth, the assumption that any illustration will facilitate comprehension without regard to the content and structure of illustration information, and similarly, the assumption that all text information will equally benefit from illustration without regard as to the content and structure of the text in which it is embedded, is replaced by a realization that each source of information may be more or less effective at conveying certain information, and that the sum total of comprehension will be affected by the degree to which text and illustration information is related to each other.

These assumption will be adhered to in the present study. Taken together, they suggest that the study of comprehension processing of illustrated text should be concerned with investigating the moment-by-moment representations constructed by individual readers as they process information from illustrations and text. In addition, comprehension should be assessed retrospectively in order to evaluate what information readers remember, the inferences they drew, and their structuring of information in long term memory. Furthermore, specific attention should be focused on the question of how readers are able to form integrated coherent representations of information from such sources.

Summary of cognitive models of discourse processing. Cognitive models and cognitive situated models of discourse processing emphasize that comprehension involves the use of interactive and constructive processes which operate at multiple levels of representation. It is currently recognized that discourse comprehension depends on a complex interplay between text structure variables, reader characteristics, and task structure and context variables. Salomon (1989) has made a similar argument for the problem of understanding how individuals learn from illustrated texts.

Extending this position to the study of illustrated text ought to be fruitful in that it helps to clarify some simplistic assumptions that have been made in the past with regard to the way in which the effects of illustrated text should be studied. In particular, such an approach emphasizes the need to assess the nature of representations that readers construct in an ongoing manner. Furthermore, this approach is helpful in that it focuses such broad questions as "Do illustrations positively affect comprehension?" to more clearly defined questions such as: "What are the cognitive consequences of providing readers with illustrated texts in terms of the kinds of representations they are able to generate as they read? What are the cognitive consequences of structuring texts and illustrations in various ways? How are cross-modal relationships signaled to the reader? How do readers integrate information within and across media?".

The methodologies and measures used to study comprehension processing have matured along with these theories. In particular, it is currently recognized that the study of discourse processing (and by extension, illustration processing) requires the use of multiple measures of comprehension that span different points in processing and which are sensitive to measuring different aspects and levels of representation.

Of particular relevance to the issue of information integration within and across media is the concept of a situation or mental model level of representation. The following section provides a review of this concept.

Mental Models

Constructing situation and mental models. According to constructivist accounts, one of the levels of representation involved in discourse comprehension is the construction of a situation model (van Dijk & Kintsch, 1983), mental model (Johnson-Laird, 1983), or conceptual frame model (Frederiksen, 1986). These terms (situation model, mental model, and conceptual frame model)[1] are used by discourse researchers to refer to a qualitatively distinct level of comprehension which represent the state of affairs described by the text rather than the content of the text itself. The impetus for proposing such a level of representation as an integral aspect of discourse comprehension stems from the observation that readers understanding and memory for discourse typically goes beyond information presented in the text.

Bransford, Barclay, and Franks (1972) provided an early demonstration of the existence of this level of representation by showing that readers' recognition memory for sentences like "Three turtles rested *on* a floating log, and a fish swam beneath *them*." was easily confused with sentences like "Three turtles rested *on* a floating log, and a fish swam beneath *it*.". However, readers did not confuse sentences like "Three turtles rested *beside* a floating log, and a fish swam beneath *them*. with sentences like "Three turtles rested *beside*

[1]The terms 'situation model', 'mental model', and 'frame' will be used interchangably and are considered by the author to be equivalent, unless otherwise noted.

a floating log, and a fish swam beneath *it*.". Although not the original intention of the authors, this finding has subsequently been interpreted as demonstrating that comprehension involves more than interpreting the propositional content of sentences. That is, readers also appear to construct a model of what the text is about. Such findings have led to the widespread acceptance of constructivist rather than interpretative accounts of discourse comprehension.

Fletcher (1994) has summarized some of the empirical findings that can be accounted for by adopting this additional level of representation, but remain difficult to explain according to a purely propositional account of comprehension. For example, the finding that individuals may differ according to their interpretation of the significance of a discourse despite similarities in comprehension of its propositional content can be explained by proposing this additional level of representation. In addition, research on human and machine translation which indicates that maintaining the propositional content of a message often does not lead to a successful translation (e.g., Hutchins, 1980) also suggests that a purely propositional account of discourse processing is insufficient or incomplete. Research has also indicated that reader's are able to appropriately and consistently reconstruct scrambled stories (e.g., Bower, Black, & Turner, 1979), thus providing evidence that readers' possess and are able to use knowledge of a canonical structure of events. Furthermore, and perhaps most convincing, research has clearly demonstrated a dissociation or distinction between "comprehension" and "learning" from discourse (e.g., van Dijk & Kintsch, 1983). The ability to account for such phenomena are at the heart of the appeal for a situation or mental model level of representation.

Features of a mental model. According to the mental model view of text processing, readers generate a high-level representation of the situation described by the text. The mental model level of representation is thought to describe the *significance* (including reference) of a discourse whereas the propositional level of representation describes the *sense* of a discourse. Johnson-Laird (1983) describes mental models as representations containing tokens corresponding to entities in the world or described in a discourse. Furthermore, the properties of these tokens and the relationships between tokens correspond to our understanding of the states of affairs that the models represent. Mental models are thought to be the result of constructive processes that *integrate* information from text with the reader's pragmatic, linguistic, and world knowledge. Thus, mental models go beyond the literal meaning of discourse and embody inferences, instantiations, and reference. It is assumed that construction takes place on-line and in parallel with the construction of a propositional text base representation. According to Johnson-Laird (1983), mental models may take the form of either a propositional network, spatial image, or temporally organized string, or some combination of these.

Mental models have been described as possessing certain attributes which enable the reader to modify and integrate information, and to reason and draw inferences from discourse. For example, Glenberg, Meyer, and Lindem (1987) have described mental models as having the following characteristics. Mental models are *updatable* and are subject to change as new information is encountered by the reader. Mental models are *manipulable*, allowing readers to perform mental simulations of entities that are dynamically related. The importance of this feature is that it enables readers to reason, particularly in a qualitative fashion, from their knowledge and may result in making some initially implicit relationships explicit to the reader. Mental models can be *perceptual-like* in that they can be used to integrate information from or about multiple sources. Mental models control inference making and influence the reader's judgment of coherence by *foregrounding* certain information or inducing a particular perspective.

Thus, mental models should enable readers to integrate disparately presented text information into a coherent structure, as well as to integrate text-derived information with their prior knowledge. That is, the construction of an effective mental model should include both types of integrative processing that were described above in the section on knowledge integration. Furthermore, in the case of dual or multimedia presentation, mental models should enable readers to integrate related information across modalities.

Mental models and the processing of spatial information from text. Most research investigating the role of situation or mental models in discourse processing has focused specifically on the processing of spatial information (e.g., Denhière & Denis, 1989; Glenberg, Meyer, & Lindem, 1987; Haenggi, Kintsch, & Gernsbacher, 1995; Morrow, Bower, & Greenspan, 1987; Perrig & Kintsch, 1985; Rinck, Williams, Bower, & Becker, 1996; Taylor & Tversky, 1992). For example, Perrig and Kintsch (1985) had subjects read either a survey or route description of an imaginary town and tested subject's memory for the actual text, for spatial inferences grounded in the text, and for maps constructed from the information in the text. The survey description presented an overview of the town and described the spatial configuration of entities using an extrinsic frame of reference (i.e., north, east, etc.). On the other hand, the route description presented information from the perspective of a specific location and described the spatial configuration of entities using an intrinsic frame of reference (i.e., to the left, above, etc.).

Perrig and Kintsch (1985) assessed the representations that readers constructed by having readers verify both locative and non locative statements related to the text as either true or false as they read. Reaction times and error rates were recorded. The authors reasoned that, if readers represent the exact wording of the text, then responses to verbatim questions should be faster and more accurate than for inference questions. On the other hand, if readers construct situation models as they read, then readers should respond to verbatim and inference questions with equal accuracy and speed. That is, the reader should not be able to distinguish statements that were actually presented in the text from statements which were inferable from the text. Furthermore, they argued that if the situation model that readers construct depends upon the particular perspective of the narrative, then readers should respond faster to inference statements framed from the perspective read than to inference and verbatim statements from the alternative perspective. If, however, the situation models represent spatial relationships independent of the perspective read, then there should be no differences on the inference questions as a result of the perspective read. Perrig and Kintsch found evidence that readers construct situation models for texts as they read. That is, readers "recognized" actually presented and inferential statements with equal accuracy and speed. In addition, performance was slightly better for route than for survey descriptions, leading the authors to conclude that there was some perspective effect.

Taylor and Tversky (1992) have criticized the Perrig and Kintsch study on the basis that the route text was more coherent and spatial relationships were more explicit than in the survey text. In their study, Taylor and Tversky investigated whether subjects who read a survey versus a route description of a fictitious geographical area constructed different situation models according to these perspectives. They employed procedures and measures similar to Perrig and Kintsch, but also attempted to balance the texts for coherence and had each subject read two texts - one from each perspective. In addition, they included a condition in which subjects studied a map. Taylor and Tversky's results confirmed Perrig and Kintsch's conclusion that readers construct situation models as they read. However, in this study no perspective differences were found. Taylor and Tversky concluded that regardless of the type of text description (survey or route) or whether subjects read a map, readers form highly similar spatial mental models to represent the salient landmarks and the spatial relationships between them.

Denhière and Denis (1989) have also investigated the issue of how different text organization structures may elicit differences in the ease with which the reader is able to generate a situation model for texts describing spatial configuration information. Specifically, these investigators compared the on-line processing (indexed by sentence reading times) and cued recall performance of subjects who received either a *linear* (i.e., scanning from left to right and from top to bottom) or a *hierarchical* (i.e., describing the configuration of entities from central point of reference) text description of the spatial configuration of six natural entities or of six unnatural entities on an imaginary island. This study is particularly interesting, since the focus is on the problem of using an inherently linear medium (i.e., text) to describe a non-linear situation (i.e., a non-linear spatial configuration). Denhière and Denis point out that the organization and sequencing of information in texts (i.e., text staging) describing spatial configuration is less constrained than for other sorts of texts. For example, in narrative texts the author's description of the sequencing of events is usually constrained by their order of occurrence. Similarly, for procedural texts the author's description of the organization of sub procedure components and alternative procedures is typically constrained by temporal and conditional relations. However, for describing the spatial configuration of objects there are few such constraints. In such a situation, the question of how to "best" organize information is an open one.

Denhière and Denis observed significantly longer reading times for sentences which violated a linear sequencing of entity configuration (i.e., the hierarchical description) and poorer graphic recall. These effects were somewhat attenuated when subjects were provided with explicit instructions to form images as they read. The authors concluded that a linear text description of spatial configuration is more compatible with the on-line elaboration of a situation model than is a hierarchical text description. Thus, in the absence of an illustration and when reader prior knowledge is low, text descriptions of spatial configurations are probably easier to comprehend when described in a linear rather than a hierarchical manner. While this study provides an effective demonstration of the role of text structure on reader's ability to generate a situation model of spatial information, it does not address the issue of how illustrated text may effect the ability to construct a situation model. It should be also be noted that the authors did not collect on-line information about the content of readers understanding. It would be interesting to investigate the extent to which such difficulties may be attenuated by the addition of an illustration. Additionally, it would be fruitful to study how different text staging structures may influence the comprehension processing of an illustrated text.

In addition to research which supports the notion that readers' construct mental models of spatial text information as they read, it should also be noted that there is some research that supports the notion that readers generate such models when processing non spatial information including, goal information (Huitema , Dopkins, Klin, & Myers, 1993), procedures (Schmalhofer & Glavanov, 1986), and linear orderings of objects (Fletcher & Chrysler, 1990).

Summary of the role of mental model in discourse processing. The various studies described above demonstrate that readers may generate high level representations of the situation described by a discourse as they read (i.e., form mental models). Mental models enable the reader to integrate related information, draw certain inferences, adopt a particular perspective, and to reason from their understanding. That is, mental models are important in supporting a variety of high level cognitive processes. While most of this research has been concerned with the processing of spatial information, there has also been some indication that readers may form mental models for non spatial information. Although, this issue requires further empirical attention.

It should also be pointed out, however, that some studies have indicated that the construction and degree of elaboration of mental models is not necessarily a routine phenomenon (e.g., Denhière & Denis, 1989; Kozma, 1991; McKoon & Ratcliff, 1992; O'Brien & Albrecht, 1992; Wilson, Rinck, McNamara, Bower, & Morrow, 1993; Zwaan & Van Oostendorp, 1993) and may vary depending upon the characteristics of the reader (e.g., goals, prior knowledge), the text organization, and the nature of the reading task. Similarly, one might also expect qualitative variations in the nature of mental models depending on the type of text information, although this issue still needs to be explicitly addressed. For example, further research is needed to address the question of how multiple types of information within a single text are integrated in mental model representations.

Mental model construction for comprehending multi-modal information. With regard to the processing of illustrated text, the construction of mental models should additionally facilitate the integration of text and illustration derived information. According to Glenberg and Langston (1992) representational elements in a mental model can index both propositional and perceptual information. It is in this way that mental models can be used to integrate information derived from separate and qualitatively different information sources. Thus, Glenberg and Langston argue that illustrations should help readers build mental models since illustrations typically depict a specific situation. For example, a mental model could be used to integrate text information describing the features of an object with information depicted in an accompanying picture which indicates the object's spatial location. Given this view, the authors have argued that illustrations can facilitate readers' construction and management of mental models in working memory. Thus, with regard to objections raised about the "routineness" of mental model construction and elaboration, one possible benefit of text illustrations is that they may alleviate some of the cognitive demands of such processing and thereby enable "routine" mental model construction.

Based on their assumptions, Glenberg and Langston tested a number of predictions. First, pictures should facilitate comprehension and retention of text. In particular, facilitation should be greatest for information that is "noticed" when a mental model is formed, but is left implicit or is difficult to understand in the text. Furthermore, pictures that encourage the noticing of inappropriate relations (i.e., conflict with the situation described by the text) may reduce comprehension and retention. Finally, these effects should be clearly attributable to a level of representation different from the representation of the text alone. In order to test their predictions, Glenberg and Langston had subjects read a series of short texts either with or without diagrams. Each of the texts described a four step procedure with the two middle steps described as co-occurring in time. The authors argue that a text-based representation of such a procedure would differ from a mental model. That is, the text presents information about steps sequentially and therefore it is expected that representational connection between the first and second step (close pairs) will be stronger than between the first and third step (far pairs). However, a mental model is able to capture the hierarchical nature of the procedure and should equally relate the first step to the second and third steps. The diagrams that accompanied texts depicted either the text-base (sequential representation of procedural steps) or the mental model for the procedure (hierarchical representation of procedural steps). The authors assessed the strength of relationships between readers representation of close and far procedural steps. Based on this data they were able to draw inferences about the representations (i.e., textbase versus mental model) that readers generated. Based on their results, Glenberg and Langston concluded that related illustrations can facilitate the construction of a mental model by increasing accessibility of (or foregrounding) certain information in working memory. The authors argue that such foregrounding induces the reader to notice certain text implicit relationships and therefore to form certain inferences as they read.

However, it should be pointed out that such foregrounding effects are not limited to illustrations but can also be elicited simply through text foregrounding (i.e. *staging*). For example, Roy (1991) studied the relationship between various text staging strategies and high and low prior knowledge readers comprehension of a routine computer procedure (i.e., moving text in a word processing document). Two text staging strategies were compared: a *hierarchical* text which presented a top-down left-right description of the component procedures and thus highlighted the hierarchical structure of the procedure, and an *enactment* text which presented a left-right bottom-up description of the same component procedures and foregrounded the linear sequence of component procedures at each level in the hierarchy. Text staging was found to exert powerful effects on the on-line comprehension processing of text as measured by reading time and concurrent verbal reports of reader's ongoing understanding of the procedure. In particular, both high and low prior knowledge groups of readers who read the hierarchical text demonstrated more intensive processing of and more complete and connected representations of the procedure than those who read the enactment text. Thus, if researchers want to claim that illustrations facilitate the construction of mental model representations in ways that are distinct from the effects of text structure (i.e., that there is something special about how text illustrations affect comprehension processing, and in particular mental model construction), they need to clarify exactly what aspects of mental model representations are affected by illustrations and how.

Hegarty and Just (1989) investigated the issue of when and why people might inspect a diagram when reading a text describing the workings of a mechanical device (e.g., a pulley). These researchers tested a preliminary model of text and diagram processing by investigating how text and diagrams are integrated during reading. Hegarty and Just assume that people may inspect diagrams for different reasons according to characteristics of the text and the reader. They describe three distinct purposes that may be served by inspecting a diagram while reading a mechanics text: (a) as an aid to the construction of a mental model of text information (i.e., mental model formation as the underlying process), (b) as a memory aid for reactivating part of a mental model derived from text information previously read (i.e., reactivation as the underlying process), and (c) to encode new information not found in the text into a mental model (i.e., elaboration as the underlying process). Furthermore, they assume that these different processes should be reflected in distinctive patterns of readers' eye fixations. For example, if a reader's diagram inspection is focused on information that has just been read, this would be interpreted as evidence for a formation process. On the other hand, if a reader's diagram inspection is focused on information that has been previously read (i.e., before the last unit of text read) this would provide evidence for a reactivation process. Finally, if a reader's diagram inspection is focused on information not provided in the text, this would be evidence for an elaboration process.

To test these hypotheses, Hegarty and Just recorded high and low prior knowledge readers' eye movements while they read one of two illustrated texts describing a pulley system. The same illustration was used in both text conditions and described the configuration of the components of the pulley. The two texts were designed to vary in the degree to which they described configuration information. One variation of the text (the longer version) provided configuration information that was redundant with information that could be obtained from the diagram, while the second variation (the shorter text) omitted such configuration information.

Hegarty and Just's found some support for formation and elaboration processes. For example, readers tended to direct their gazes towards relevant components of the diagram after encountering a text description of that component (i.e., formation inspections). This was especially true for low prior knowledge readers who received the longer version of the text. In addition, readers tended to direct their gaze towards information provided in the diagram that was not described in the text (i.e., elaboration inspections). This was especially characteristic of high prior knowledge readers who received the shorter version of the text. This pattern of results was anticipated by Hegarty and Just who reasoned that low prior knowledge readers would consult diagrams when trying to comprehend configuration information described in the text. High prior knowledge learners, on the other hand, would only need to consult a single source to gain such information. Apart from these anticipated results, however, readers allocated the majority of their diagram inspection time towards components they had previously read about or inspected. While it was hypothesized that this pattern would be evidence for the process of reactivating previously comprehended information, Hegarty and Just report that the frequency and duration of inspections suggest instead that readers were actually attempting to gain new information and to integrate disparately presented text information. Based on this finding, they added a fourth possible purpose for diagram inspection - as a context for integrating disparately presented information from the text.

As Hegarty and Just point out, these results should be interpreted as preliminary since their study failed to employ any measure of comprehension of the text or of the diagram. Thus, little can be said about whether or how the observed patterns of eye fixations are related to differences in comprehension.

In a follow up study, Hegarty and Just (1993) have extended their work to include measures of comprehension on a larger sample size of readers. These studies were reviewed in detail above in the section on the integration of text and illustration information. However, for the sake of clarity, a brief repetition of their conclusions is warranted here. Hegarty and Just concluded that: (a) readers need to process both media to form a complete mental model, (b) readers integrate information from text and diagrams incrementally from a local representation of several related components to a global representation of the entire system, and (c) low mechanical ability readers experience more difficulty constructing a mental model representation from illustrated text and require longer study times, and more frequent text regressions and diagram inspections than their high mechanical ability counterparts.

Both the 1989 and the 1993 studies are limited, however, in that eye fixation data are limited to indexing what is being looked at, when, by the reader, and does not in itself provide information with regard to the content of representations generated by comprehension processes. It could be argued that a more informative approach to the study of comprehension processing with regard to illustrated text would involve both on-line and off-line measures of the content of readers' comprehension.

Over the past 20 years Mayer and his colleagues have engaged in extensive research concerned with investigating the instructional effects of illustrations (see Mayer, 1993; 1997 for reviews of this work). His recent works have focused more specifically on how explanative illustrations may help readers (particularly low prior-knowledge readers) build runnable mental models of cause-effect systems. Mayer and his colleagues have proposed that illustrations which provide information about (a) system topology of the device described (i.e., component parts and their relationships), and (b) component behavior of the device (i.e., states and changes in states) in a spatially contiguous manner (i.e., the contiguity principle) facilitate learning by helping readers to build referential and associative connections between verbally and visually presented information.

The contiguity principle states that more meaningful (i.e., flexible) learning results when cross-modal information is presented in a spatially/temporally contiguous manner than when presented in isolation. Mayer proposes that this principle rests on assumptions from dual-coding theory which posits two separate but interconnected processing systems - one for representing verbal information, and another for representing visual or image information (Paivio, 1986; 1991). The contiguity principle specifically addresses the issue of building interconnected representations for which the learner must build *referential* connections between the two processing systems. Thus, the contiguity principle predicts that the learner will build more referential connections between visual and verbal information when cross-modal information is presented contiguously than when presented distally.

Mayer and Gallini (1990) have tested and confirmed several hypotheses with regard to the ability of illustrations to facilitate mental model construction and with regard to the contiguity principle, including the following: (a) explanative illustrations will lead to increased recall of explanative but not non-explanative information, (b) explanative illustrations will lead to an increase in creative problem-solving (i.e., evidence of a mental model representation) but not verbatim retention (i.e., a surface level representation), (c) explanative illustrations will increase conceptual recall relative to non-explanative illustrations, (d) explanative illustrations will lead to increased problem-solving relative to non-explanative illustrations, (e) explanative illustrations will lead to increased recall for low prior knowledge individuals but not necessarily for high prior knowledge readers, and (f) explanative illustrations will increase problem-solving for low prior knowledge readers but not for high prior knowledge readers. While these authors appeal to dual coding theory to explain their finding, they also concede that a discourse situation model perspective (e.g., Kintsch, 1989) provides equivalent explanatory power.

Mayer has also attempted to extend his finding on the contiguity principle from research on learning from illustrated text to learning from animation (e.g., Mayer & Anderson 1991; 1992, Mayer & Sims, 1993) For example, Mayer and Anderson (1991) report two experiments which were designed to examine the instructional efficacy of five conditions of structuring animation and narration sequences in computer-based instruction. Experiment 1 was designed to compare the problem solving and verbal retention performance of students who received concurrent (i.e., temporally contiguous) versus successive (i.e., temporally non-contiguous) presentation of animation and narration sequences describing how a pump works. The learning performance of eight groups of university students was compared: (a) group 1 received concurrent animation and narration; (b) groups 2, 3, 4, and 5 received successive animation, however the order of animation and narration was varied across groups; (c) group 6 received animation only; (d) group 7 received narration only; and (e) group 8 received no instruction (i.e., the control group). For each condition the stimulus was repeated three times to ensure that learners had enough time to process the material as the rate of presentation was not under the learners control. The group that received concurrent animation and narration demonstrated superior performance on measures related to mental model construction than did all other groups. In addition, with the exception of the control group, these groups did not differ in terms of their level of literal comprehension suggesting that, like illustrations, animation specifically facilitates high level comprehension processing when the two modes of information are physically (in this case temporally) contiguous.

While Mayer has consistently concluded that his results support the contiguity principle and the dual coding model of memory, he has also suggested that his results can be explained by Kintsch's notion of a situation model. Furthermore, the notion that the contiguity principle is particular to the problem of cross modal integration is questionable. Prior research has demonstrated that temporal/spatial contiguity is a *generally* important characteristic for integrating knowledge. That is, regardless of modality, integration is less likely to occur when related material is presented distally than when it is presented proximally (e.g., Hayes-Roth, 19 ; Sweller, Chandler, Tierney, & Cooper, 1990). Thus, Mayer has yet to show how the principle of contiguity is specific to relating cross-modal information (visual and verbal), and thus how it is necessarily derived from dual-coding theory.

In addition, while Mayer reports that the facilitory effects of contiguous presentation of illustrations and text are specific to low prior knowledge learners, a necessary condition for such a beneficial effect is an adequate level of picture-reading skill and prior knowledge in order to extract, understand, and integrate the relevant information from the presented illustration (Peeck, 1993). Such abilities also need to be more systematically investigated. That is, how is it that low prior knowledge readers are able to do this?

Summary of mental models and comprehension of illustrated text. Research has supported the notion that adding illustrations and/or animation to text specifically affects comprehension processing at the mental model level of representation (Glenberg & Langston, 1992; Hegarty & Just, 1989; 1993; Mayer & Anderson 1991; 1992, Mayer & Gallini,1990; Mayer & Sims, 1993). One explanation for this effect is that illustrations provide the reader with an external situation model of the referents of discourse content thereby providing the reader with a contextual structure in which to integrate various relevant information. This idea fits well within the situated discourse framework described by Frederiksen and Donin (1996) that was reviewed above in the section on cognitive models of discourse processing. However, studies which explicitly demonstrate that this is how readers are actually using depicted information has yet to be conducted. Furthermore, studies which have focused on how illustrations may provide a context for the interpretation and integration of text information have failed to address the question of how text affects illustration processing.

This research has also demonstrated that the degree and nature of facilitation depends on the characteristics of the reader, the text, the illustration, and the task. Based on past research, it appears that the presence of illustrations is most likely to affect mental model construction when (a) the readers is relatively unfamiliar with the content domain, (b) the text describes a dynamic functional system with spatial information (c) the illustration depicts spatial configuration information and (d) the reading task requires the reader to form a high level representation of the material.

Aside from assessing the presence of mental models, some of this research has also been concerned with indexing the on-line construction processes involved in comprehending illustrated text and in building mental models (e.g., Hegarty & Just, 1989; 1993). However, such measures have been limited to the collection of eye movement data, and as such do not provide any direct information with regard to content of readers' on-going comprehension processing. Thus, although the theoretical concept of a mental model seems promising for describing how readers process illustrated text, we currently have an incomplete description of exactly what readers are doing and understanding as they read. Further research is needed to address this issue.

The purpose of the present study is to investigate the semantic representations that readers generate as they read text , illustrations, and illustrated text from the perspective of current models of cognitive discourse processing and cognitive situated discourse. Of particular interest is the question of how readers form coherent integrated representations (i.e., mental models) as they read such materials.

Chapter III: Rationale

The research literature reviewed in the previous chapter provides clear support for the notion that adding relevant illustrations to text can significantly facilitate readers' ability to comprehend and remember information described by both media (e.g., Haring & Fry, 1979; Lesgold, 1978; Levie & Lentz, 1982; Moore & Readence, 1981; Schallert, 1980). In addition, there is equivocal support for the hypothesis that illustrations may also facilitate the processing of non-illustrated information described by the text (e.g., Levie & Lentz, 1982). However, because of conflicting results, this issue requires further investigation before any specific conclusions can be made.

By and large, previous studies have focused on measuring the products of comprehension (i.e., recall, recognition) rather than the processes involved in generating representations of the information described by text and accompanying illustrations. As such, it is not possible to gain insight into how illustrations specifically affect comprehension processing, and vice versa. Nor, is it possible to adequately evaluate the many different theoretical hypotheses that have been proposed to explain this effect.

Some of the problems with such studies is that they are often based on simplistic assumptions including: (a) the notion that illustrated information is "transparent" or "easy" to understand, (b) the assumption that memory for text and illustrations is equivalent to comprehension and learning, (c) the idea that the problem of how illustrations affect text processing can be adequately studied without considering how characteristics of the text may affect illustration processing and vice versa, (d) the assumption that individual differences in how readers process illustration information are not of interest, (e) the assumption that ease of comprehension steadily increases with number of representations provided to the reader, and (f) the assumption that any illustration will facilitate comprehension without regard to the content and structure of illustration information, and similarly, the assumption that all text information will equally benefit from illustration without regard as to the content and structure of the text in which an illustration is embedded. In the previous chapter, it was pointed out that such assumptions could be replaced by a more psychologically realistic and integrated set of assumptions within the framework of cognitive models of discourse processing and situated discourse processing. Essentially, such models assert that comprehension is a highly constructive and inferential process involving multiple levels of representation which ultimately depends on a complex interplay between the reader, the stimulus materials, the task, and the reading context.

One intriguing and currently favored theoretical explanation derived from cognitive theories of discourse processing is that illustrations help readers generate mental model representations of text information. In particular, it has been suggested that because illustrations provide an external model of the discourse situation, their presence provides the reader with a particular context in which to interpret and reason about the text. As such, the presence of text illustrations should support the reader's ability to integrate information from the text as well as to integrate text-derived information with their prior knowledge.

Studies investigating the effects of illustrated text have concluded that the degree to which the construction of mental models is facilitated may depend on such factors as (a) the type of text information (i.e., spatial information appears to be particularly conducive to mental model construction, although there is also some support for the notion that illustrations facilitate the comprehension of non spatial information), (b) the prior knowledge of the reader (i.e., low prior knowledge readers appear to be more likely to benefit from illustrations than high prior knowledge readers, (e.g., Mayer; 1993a; Hegarty & Just, 1993), (c) the contiguity between text and illustration information (i.e., the more physically integrated the two modes, the more likely readers are to generate an appropriate mental model representation (e.g., Mayer, 1993a; Sweller, Chandler, Tierney, & Cooper, 1990), and (d) the nature of the reading task (i.e., the probability that readers will generate a mental model representation is related to the reader's goals, e.g., Denhière & Denis, 1989; McKoon & Ratcliff, 1992).

Some of the mental model studies have been concerned with indexing readers' on-line processing of text and illustration information. However, no research has been conducted which has adequately attempted to measure the content of readers' on-line comprehension processing. In addition, with the notable exception of Mayer's work, few studies have attempted to measure the effects of illustrations on both the text based and mental model levels of representation. Thus, there is a specific need to carry out such research.

In addition to the issue of how illustrated texts affect the representations that readers generate as they read, some researchers have attempted to describe the different text and illustration access patterns that readers use as they process illustrated text. Again, however, such studies can be criticized on the basis that they have not attempted to measure the effects of such differences on comprehension processing per se. Thus, there is a specific need to more clearly assess natural variation among readers' use of illustrated text and the consequences that such patterns of use have on comprehension processing.

The Current Study

The purpose of the present study is to begin to address such shortcomings by investigating readers' on-line comprehension processing of illustrated text. The study focuses on the nature of semantic representations that low prior knowledge readers are able to construct as they read from text and illustrations which present multiple types of information describing a functional system (i.e., the human visual system). The study focuses on low prior knowledge readers since past research has demonstrated that such readers are more likely to benefit from illustrations (e.g., Hegarty & Just, 1993; Mayer, 1993a). Indeed, it is with low prior knowledge readers that one would expect to most clearly have access to the processes of meaning construction, since high prior knowledge readers may well be able to bypass such processing. Text and illustrations that describe multiple types of information were chosen to investigate the issue of how readers integrate various types of information both within and across media. Text and illustrations which describe a functional system were chosen since past research has demonstrated that such materials are conducive to mental model building and assessment (e.g., Chi et al., 1994; Mayer, 1993a; Hegarty & Just, 1993).

Of particular interest is the issue of how readers use and integrate these sources of information to construct mental model representations as they read. How does each source of information contribute to the reader's understanding? How does one source of information influence the manner in which the reader processes additional sources of information? What kinds of inferences are supported by each source of information? What kinds of integration are supported? The current experiment was designed to answer these questions and includes the following conditions: (a) a *text only* condition (TO), (b) an *illustrations only* condition (IO), (c) a *text and illustration* combination where illustration access is experimentally controlled (CA), and (d) a *text and illustration free access* combination condition where the reader's access to text and illustrations is unlimited (i.e., the most natural reading situation) (FA). Pre-planned comparisons among these groups will address the issue of the individual and reciprocal effects of text and illustrations on comprehension processing under controlled and natural reading conditions.

In addition, a second issue investigated in this study is the question of how the processing benefits of illustrated text may depend on the particular perspective from which the target domain is described (i.e., information staging). To investigate this issue the experimental text and illustrations were constructed to highlight one type of information (i.e., *structure* information) about the human visual system over other types of information (*function* and *energy* information). Pre-planned comparisons between the comprehension processing of the different information types were made to investigate the question of how aspects of mental model construction depend on the type of information perspective highlighted and how such effects may interact with different patterns of text and illustration use.

Investigation of these issues requires the use of multiple indices of comprehension processing, including both on-line and off-line measures. On-line measure of sentence processing times, text look-backs, and illustration access times were used to index where reader's devoted "heavy" processing efforts during information comprehension. These data were supplemented with an on-line interpretation task in which readers provided an ongoing account of the content of their understanding at both literal and inferential levels. In addition, a trace of where individuals accessed illustration and text information in the "free access" condition was collected to provide data about how readers "naturally" used illustrations and text. Post input verbal and visual recalls were collected to provide data about what information readers understood and were able to recall or infer from the materials, how they organized and integrated that information, and how they were able to construct mental models of each information type. Finally, a set of post-input questions was used to assess readers' ability to integrate domain information and to use their current understanding to generate new information.

Specific Research Questions & Hypotheses

The effect of type of illustration access to in the context of text. The first research question investigated concerns potential differences between the two conditions in which participants were exposed to a combination of text and illustrations (FA and CA). This was examined in a pre-planned contrast between the *free* access and *controlled* access conditions. In particular, this comparison focused on assessing the effect of forcing a particular coordinated pattern of access between media (CA) on processing and learning versus allowing readers to self-determine when to consult illustration information (FA) in the context of text. The primary purpose of including this comparison was to be able to detect any oddities in the experimentally controlled access to illustrations condition (CA). That is, while the CA condition is necessary to this experiment in that it ensures a consistent pattern of illustration access across participants in that group, it is also of concern that the manner in which this is achieved does not depart substantially from the way in which readers would normally use illustration information.

On a secondary note, this contrast was also of interest in addressing the question of whether individual differences in patterns of coordination between media are associated with different comprehension processing patterns and learning outcomes. Based on some past research (e.g., Mayer, 1994; Chandler & Sweller, 1991) one might expect that these two experimental groups would differ from one another in the mental models they are able to construct. Specifically, the reader's pattern of movement between text and illustration sources might affect the reader's perception of cross modal cohesion and thus affect the nature of mental models that will be constructed. However, the particular characteristics of such differences cannot be specified *a priori* since in the free access condition movement between media is under the reader's control.

It should be noted that if significant differences were found between the CA and FA conditions, then these two conditions would be separately compared to the TO and IO conditions to test the hypotheses listed below. If, however, the CA and FA conditions are found to be highly similar, then they will pooled for the purposes of testing the subsequent hypotheses.

<u>The effect of illustrations on text processing</u>. The second research question addressed in the current study concerned the effect of exposure to illustrations on text processing. This question was addressed in a pre-planned contrast comparing the measures of comprehension and processing of the two text with illustrations groups (FA and CA) to those of the text only group (TO).

Based on the findings of past research, it was expected that exposure to a combination of text and illustrations (FA and CA) would facilitate low prior knowledge readers' ability to construct a coherent integrated representation (i.e., a mental model) of the target domain. Specifically, it was expected that the presence of relevant illustration information in the context of text would support the on-line construction and elaboration of mental model representations by providing a constrained environment in which to understand and reason about text information.

In the context of the current experiment this hypothesis implies that when compared to the text only group (TO) the two groups exposed to a combination of text and illustrations (FA and CA) should demonstrate evidence of: (a) more information integration in their on-line protocols, (b) more elaborated and integrated mental model representations of the human visual system, and (c) superior ability to answer comprehension questions that require information integration and the generation of new knowledge.

No specific pattern of differences between the groups exposed to text with illustrations and the text only group were predicted on the remaining comprehension measures (i.e., processing time, verbal recall, and visual recall). However, there are several possible outcomes. In terms of sentence processing time, one possibility is that the addition of illustrations to text information (FA and CA) will be associated with greater cognitive demands as compared to the text only group (TO), resulting in the need for longer processing times in comprehending text propositions. This bottom-up account of multimedia processing follows from the view that readers in this situation are required to comprehend information in each medium separately, and also to integrate information across modalities. On the other hand, it is also possible that the CA and FA groups will demonstrate faster sentence processing times. This pattern of results would be consistent with a more top-down view in which the presence of illustration information facilitates or supplants some aspects of comprehension processing at a conceptual level that would otherwise be necessary to perform on the basis of text alone. Finally, if no differences in overall processing time were found this would support the position that the cognitive demands of the two types of learning situations do not differ substantially from one another.

The effect of text on illustration processing. The third research question concerned the effect of exposure to text on illustration processing. This issue was addressed in a pre-planned contrast comparing measures of comprehension and processing of the two text with illustrations groups (FA and CA) to that of the illustrations only group (IO).

It was anticipated that the presence of text would serve an elaboration function by providing readers with semantic information necessary to successfully interpret and integrate the information presented through the illustrations. Specifically, this hypothesis implies that when compared to the illustrations only group (IO), the two text with illustrations groups (FA and CA) will demonstrate: (a) more on-line elaboration and integrative processing, (b) more complete visual and verbal recalls, (c) more elaborated and integrated mental model representations, and (d) superior ability to answer comprehension questions that require information integration and the generation of new knowledge.

The effect of information perspective. Finally, it was expected that the comprehension processing of readers in all experimental groups would demonstrate a bias towards *structure* information (i.e., the conceptual perspective that is emphasized in the materials) over *function* and *energy* information. This hypothesis was tested in a pre-planned contrast in which the measures of comprehension and processing of structure information were compared to those for function and energy information. The rationale underlying this hypothesis is a logical extension of past research documenting the effect of information staging on text processing (e.g., Britton, Glynn, Meyer, & Penland, 1982; Clements, 1979; Kieras, 1981; Marshall & Glock, 1978) in which similar effects on the comprehension processing of illustrations and text were expected. From a theoretical point of view, this hypothesis implies that information staging is an important dimension of both text and illustrations and this dimension significantly influences comprehension processing for both modes of information representation.

In the context of the current experiment, evidence of information perspective effects should be manifest in a pattern of (a) longer processing times for structure information than for function and energy information, (b) proportionally more on-line interpretations describing structure information, (c) greater proportional recall of this information in both verbal and visual modalities, and (d) more elaborated and connected mental models of structure information.

Finally, it was expected that the effects of information perspective would constrain the benefits of adding illustrations to text specifically to the comprehension processing of structure information. That is, the benefits associated with exposure to a combination of text and illustrations are expected to be larger for processing and comprehension of structure information than for energy path or function information.

Chapter IV: Method

The current study was designed to contribute to a better understanding of the individual and reciprocal effects of text and illustrations as information sources on comprehension processing and readers' ability to comprehend and integrate information both within and across media. As such, this experiment involved varying the sources of information to which readers have access (i.e., text, illustrations, or a combination of text and illustrations). Past research has demonstrated that, in general, low prior knowledge readers benefit from the addition of relevant illustrations to text. Precisely how such readers use, understand, and integrate information from these sources however is not well understood.

The current study focused on investigating the nature of semantic representations that low prior knowledge readers are able to construct as they read from text, illustrations, and a combination of text and illustrations which present multiple types of information describing a functional system (i.e., the human visual system).

A second, main objective of this study was to determine whether the effect of information perspective (i.e., staging) on readers' ability to comprehend and integrate information from text generalizes to text and illustrations and to illustrations alone. Past research on information staging has been limited to prose and has not adequately investigated the question of how such effects may be different for comprehending illustrated text or for illustrations alone.

Participants

Twenty-four volunteer undergraduate students were paid participants in this study. All participants were English first-language speakers. Participants were drawn from the domain of mechanical engineering. This particular population was chosen since these students are not required to take any course(s) that may deal with the content domain used in the present study and are therefore appropriate candidates for assessing how readers comprehend and learn novel information. Furthermore, this population of students do frequently encounter the kind of text structure and illustration types used in the present study (i.e., descriptions of functional systems). Thus, the selection procedure was designed to avoid the possibility of confounding prior knowledge of content with prior knowledge of text structure and illustration type. Six participants were randomly assigned to each of the four experimental conditions.

Sample Characteristics

Pre- and post-task questionnaires were employed in order to establish that the content of the stimulus materials was appropriate for the participants sampled. The pre-task questionnaire was designed to assess participant's prior domain knowledge before they were exposed to the experimental materials. The questionnaire was administered to participants at the beginning of the experimental session. Participants were asked to: (a) list all academic courses previously taken that dealt with the topic of the human visual system, (b) list the type and level of course in which this content was encountered, (c) and list the year the course was taken, and the number of classes devoted to the topic. In addition, participants were asked to write a brief paragraph describing their current understanding of the human visual system (see Appendix A).

The post-task questionnaire was administered to participants at the end of the experimental session to obtain general information about readers' experience with the domain content, and the text structure and illustration types used in the present study. The questionnaire contained several items in which participants were asked to rate the following items of a 7 point Likert scale: (a) their familiarity with the content of the experimental materials, (b) the difficulty of the text, (c) the difficulty of the illustrations, (d) the frequency of occurrence of similar diagrams in their area of study, (e) the frequency with which they use such diagrams in their area of study, and (f) the difficulty of such diagrams in their area of study. In addition, several free response items were included (see Appendix B).

Participants' responses to the pre-task and post-task questionnaire clearly established that they were relatively unfamiliar with the content of the stimulus text and illustrations. An examination of responses to the pre-task questionnaire revealed that participants had little formal experience with the domain of the human visual system and expressed limited prior knowledge related to a very small subset of the concepts covered by the experimental materials (e.g., lens, iris, pupil). Table 1 below provides the averages by condition and across condition for participants' responses to items on the post-task questionnaire. Each of these items was rated on a 7 point Likert scale with 1 indicating absence of the rated quality, 4 indicating neutrality, and 7 indicating complete presence of the rated quality. Thus, the participants in this study rated themselves as relatively unfamiliar with the content of the materials to which they were exposed ($M = 3.94$). Furthermore, they found the materials to be slightly easy ($M = 5.06$), rated the illustrations used as helpful (M = 5.80), and rated the illustration types to be similar to those encountered in their own area of study (M = 6.38). Therefore, the stimulus materials appeared to be suited to the purposes of the present study in that they presented novel information that was somewhat challenging, but used familiar means to do so.

Table 1

Average Post-task Questionnaire Responses by Presentation Condition

	Presentation Condition				
	FA	CA	TO	IO	Average
Question	M	M	M	M	M
Prior knowledge of content	4.83	3.33	3.50	4.08	3.94
Ease of text	5.00	5.17	5.00	X	5.06
Aide provided by illustrations	6.17	5.83	X	5.40	5.80
Familiarity with illustration type	6.67	6.00	5.20	5.00	5.72
Typicality of illustration type in area of study	6.83	6.67	5.83	6.17	6.38
Ease of understanding illustration type in area of study	6.17	6.00	5.33	5.08	5.65

Note. All questions employed a 7-point scale. FA = free access condition; CA = controlled access condition; TO = text only condition: IO= illustrations only condition; X = question not applicable to condition.

One characteristic of the sample used in the current study that was explicitly measured was spatial ability. The Paper Folding and Surface Development tests (Ekstrom, French, Harman, & Dermen, 1976) were administered to each participant. Participants were uniformly high in terms of spatial ability as measured by these tasks (see Table 2 below).

A one-way MANOVA was performed on participant's' spatial ability scores and revealed that the experimental groups did *not* significantly differ from one another on these dimensions (multivariate F (6,38) = 1.344, p = .262). Pearson correlations also revealed that spatial ability did not significantly correlate with any of the dependent measures of interest (see Appendix C). This result may be due to the fact that participants were quite uniform on this measure. Although, spatial ability may indeed significantly predict performance on learning from illustrated materials for populations who vary more dramatically on this ability, for the purposes of the present study spatial ability was not regarded as an important moderating variable and will henceforth be ignored.

Table 2

Average Spatial Ability Scores by Presentation Condition

	Paper Folding		Surface Development	
Condition	M	SD	M	SD
Text Only	8.167	0.753	27.833	1.472
Illustrations Only	7.667	1.966	21.333	4.502
Free Access	7.833	1.329	25.500	5.320
Controlled Access	8.333	2.066	26.000	3.225
Total	8.000	1.532	25.167	4.380

Materials

The domain. The human visual system was selected as the domain of investigation. In addition to meeting the considerations described in the rationale section (i.e., that the domain be conducive to the construction and assessment of mental models), the choice of the domain was made on the following grounds. First, comprehension of how the human visual system works requires that the learner understand and integrate several types of information including: (a) a description of the component parts of the eye and their configuration within the system (i.e., *structure* information), (b) a description of the various behaviors and functions carried out by each part or subset of parts of the system (i.e. *function* information), and (c) a description of the flow and transformation of energy at various points in the system (i.e., *energy* information). Second, investigation of numerous books presenting this domain revealed that illustrations were commonly used as adjuncts to textual descriptions of the human visual system. Interestingly, the kinds of diagrams used were highly consistent across texts (typically one describing the components of the eye with a separate diagram depicting a blowup of the retinal structure). Thus, this domain is appropriate to the study of how readers integrate information both within and between media. Furthermore, because the domain involves more than one type of information, it lends itself to the control and study of the effects of a particular pattern of text perspective (i.e., *staging*) which may privilege one type of information over others. The effects of information perspective on the comprehension processing of text and illustrations have not been previously studied.

Original illustration characteristics. Two basic illustrations were selected from existing textbooks and subsequently modified to be used in the present study. They were originally gray scale schematic pictures depicting structural and configural information about the human eye. Both illustrations were scanned into a computer format from original textbook illustrations for the purpose of computer presentation. These illustrations were selected as *typical* examples of text illustrations used to describe information in this content domain. In addition, both were naturally occurring text illustrations designed for undergraduates and provide clear schematic depictions of important structural information.

The computer display of illustration 1 was 8.333 by 4.722 inches (600 by 340 pixels) in size. This illustration depicts a cross-sectional view of the human eye and includes labeled lines, which point to and name various components of the eye. The source for this illustration is E. B. Goldstein's 1984 book *Sensation and Perception (Second Edition).* According to either Levin's (1981) or Mayer's (1993a) taxonomy, this illustration would be classified as serving an organizational function. That is, the illustration depicts relations among a set of component elements and thereby might theoretically be expected to facilitate the processes of selection and organization of information according to the elements and relations depicted.

The computer display for the second illustration was 4.861 by 5.417 inches (350 by 390 pixels) in size. This illustration represented a blown up or enlarged view of the various cell structures found in the human retina and their relative location. Illustration 2 contains alphabetically labeled structures that index the various types of cells found in the human retina with a key. The illustration could be classified as organizational according to existing taxonomies of illustrations.

Modified illustration characteristics The two original illustrations were modified in two ways for the purposes of this study. First, in an attempt to more evenly balance the type of information available from both the text and the illustrations, energy path information was added to the two illustrations. For illustration 1 (cross section of the human eye), this modification involved the superimposition of lines depicting the shape and direction of energy flow through the various structures of the eye. For illustration 2 (retinal cells), arrows depicting the direction of energy flow were added to the diagram. Although *not* directly depicted in the illustrations, it is possible to *infer* functional information from the depiction of energy path information. For example, the depiction of a change in the shape of light as it flows through a component implies that the component is involved in admitting and refracting light. No attempt was made to directly include functional information in the illustrations since this would have involved either adding animation, substantial textual labeling, or adding a substantial amount of further detail. Such modifications or additions are beyond the scope of the present study.

The second way in which illustration information was modified involved control of the perspective from which information was described. This modification was achieved in two ways. First, color was added to the particular type of information to be highlighted (e.g., structural) and information not being highlighted (e.g., energy path) was dimmed. Second, four versions of each illustration were constructed in which the content of the type of information being highlighted was sequentially built up in a manner coinciding with the textual description of information. That is, relevant structural information about a particular component was dynamically introduced into the illustration as it is described in the text. The resulting eight illustrations (i.e., four sequenced versions of illustration 1 which highlight *structural* information, and four sequenced versions of illustration 2 which highlight *structural* information) are available in Appendix D.

Original text. A number of texts describing the anatomy of the visual system were consulted in an attempt to create a composite text which incorporated typical features used to describe the domain of the human visual system. Text descriptions were quite consistent across text sources and typically mixed several types of information (e.g., component parts and their configuration, physical attributes, function of components, and a description of the light path and transduction of light energy into neural energy).

Text characteristics. The text that was created and used in the present study is a nine paragraph expository text entitled 'The Human Visual System'. This text interweaves a description of the structure of the human eye, with a characterization of the functions of each component, and the path that light energy takes as it passes through these components and is eventually converted from photon energy into neural energy. The total text is 71 sentences long, with a total of 1141 words. The average number of words per sentence is 16.07. The average number of propositions per sentence is 4.48.

The bulk of this text was adapted from R. N. Haber and M. Hershenson's (1973) book entitled *The psychology of visual perception* and supplemented with information from E. B. Goldstein's (1984) book entitled *Sensation and perception* and Keeton's (1976) book entitled *Biological science*. The extract from Haber and Hershenson's book was selected since it provided a clear and concise introductory description of the human visual system at a level appropriate to undergraduates. The intention was to use as natural a text as possible, but to provide low prior knowledge readers with enough detail to understand the anatomy and functioning of the human visual system.

Control of text perspective on domain knowledge (information staging). Since the text presents multiple information types, it was possible to control the text organization according to information type and to study the effects of the controlled conceptual perspective of the text on comprehension processing. The text structure was manipulated to highlight *structural* information. That is, information was ordered and introduced (i.e., *staged*) within sentences and paragraphs so that the text privileged *structural* information over functional and energy information. Basically this was accomplished by manipulating the topic-comment structure of sentences so that functional and energy information was always introduced in the context of structural information. In addition, the first sentence in each new paragraph (i.e., the topic sentence for the paragraph) described structural information. This structurally staged text served as the text stimulus for the experiment (a complete version of the text is available in Appendix E).

Analysis of text information. A propositional analysis was performed on the stimulus text using to Frederiksen's (1975; 1986) system. This procedure yields a literal semantic analysis of the text content. The propositions were then categorized by type (i.e., event, state, system, and relation). The breakdown of the text by proposition types is listed in Table 3 (see Appendix F for the complete propositional analysis).

Table 3

Frequency and Percent of Proposition Types

	Proposition Type				
	Event	State	System	Relation	Total
Frequency	89	109	33	87	318
Percent	27.99	34.28	10.38	27.36	100

Each text sentence was also classified according to the type of information it described. Three types of information were distinguished: (a) structure, (b) function, and (c) energy path. Structural information describes the component parts of the visual system, their number, and relative location (i.e., the parts and configuration of the system). Functional information describes the various processes and operations carried out by the component parts of the system that enable it to function as a whole. Energy path information describes the state, location changes, and direction of the energy flowing through the eye and includes description of both light energy (photon energy) and neural energy (electrical energy). An example of each information type is presented in Table 4. Most of the text sentences could be characterized as structural or functional, with fewer sentences describing energy path information. In addition, several sentences provided more than one kind of information. The breakdown of text sentences according to information type is presented in Table 5 (coding of each sentence by information type is also available in Appendix F).

Table 4

An Example of Each Text Information Type

Type	Example Sentence
Structure	The outside opening at the front of the eye is covered by a clear membrane called the cornea.
Function	The shape of the cornea is responsible for about 70 percent of the eye's focusing power.
Energy	Light first passes through this structure on its way to the retina.

Table 5

Frequency and Percent of Text Sentences Classified by Information Type

	Information Type				
	Structure	Function	Energy Path	Mix/Other	Total
Frequency	36	22	9	4	71
Percent	50.70	31.00	12.68	5.63	100

Next, each proposition was cross classified by information type (structure, function, or energy path) and proposition type (event, state, system, relation). This classification enabled the researcher to investigate how readers' on-line processing and recall is related to the semantic structure and content of information for the different reading conditions. Table 6 shows the frequency and column percents for propositions classified by information and proposition type. Table 7 shows the breakdown of propositional relation types classified by information type. Table 8 presents the average number of propositions per sentence for each information type.

Table 6

Frequency and Column Percent of Proposition Type by Information Type

Proposition Type	Structure		Function		Energy Path		Mix		Total	
	N	%	N	%	N	%	N	%	N	%
Event	26	19.12	42	32.56	19	45.24	2	18.18	89	27.99
State	63	46.32	29	22.48	14	33.33	3	27.27	109	34.77
System	12	8.24	21	16.28	0	0	0	0	33	10.38
Relation	35	25.74	37	28.68	9	21.43	6	54.55	87	27.36
Total	136		129		42		11		318	

Note: "Information Type" is printed as a spanning heading over the Structure, Function, Energy Path, Mix and Total columns.

Table 7

Frequency and Column Percent of Relation Type by Information Type

Relation Type	Structure		Function		Energy Path		Mix		Total	
	N	%	N	%	N	%	N	%	N	%
And	7	20.0	4	10.8	1	11.1	1	16.7	13	14.5
Category	1	2.9	0	0.0	0	0.0	0	0.0	1	1.2
Causal	1	2.9	3	8.1	1	11.1	3	50.0	8	9.2
Conditional	2	5.7	13	35.1	0	0.0	1	16.7	16	18.4
Equivalence	7	20.0	3	8.1	0	0.0	0	0.0	10	11.5
Equiv:Temporal	0	0.0	3	8.1	0	0.0	0	0.0	3	3.5
Identity	2	5.7	3	8.1	0	0.0	0	0.0	5	5.8
Order:Location	0	0.0	1	2.7	3	33.3	0	0.0	4	4.6
Order:Temporal	1	2.9	1	2.7	4	44.4	0	0.0	6	6.7
Possessive	2	5.7	1	2.7	0	0.0	1	16.7	4	4.6
Prox.: Attribute	1	2.9	0	0.0	0	0.0	0	0.0	1	1.2
Prox.: Identity	1	2.9	0	0.0	0	0.0	0	0.0	1	1.2
Part	10	28.6	5	13.5	0	0.0	0	0.0	15	17.2
Total	35		37		9		6		87	

Note. Equiv:Temporal = temporal equivalence ; Prox.: Attribute = proximal attribute; Prox.: Identity = proximal identity.

Table 8

Average Number of Propositions per Sentence by Information Type

	Structure	Function	Energy Path
Average	4.00	5.86	4.67

In addition to proposition and sentence information type (available in Appendix F), a conceptual model of each information type is presented in Appendix G. These models were used to assess each reader's ability to generate high level conceptual representations (i.e., mental models) of each information type, as well as the degree to which they were able to integrate this information.

Design

The experimental design consisted of four presentation conditions (illustrations only, text only, free access to text and illustrations, and text with controlled access to illustrations) crossed with three information types (structure, function, and energy information).

Participants were randomly assigned to one of four presentation conditions for the structurally staged materials: (a) an *illustrations-only* condition (IO) to be used to a baseline for assessing the comprehension processing of illustration information, (b) a *text-only* condition (TO) to be used as a baseline for indexing the comprehension processing of text information, (c) a *free access to text with illustrations* condition (FA) which represents the most natural reading situation involving illustrated text, and (d) a *text with experimentally controlled access to illustrations* condition (CA) in which assess to structural illustrations was controlled by the experimenter to be contingent upon structural text information. The stimulus materials across all four conditions involve text and/or illustrations in which structural information is staged high. It should also be noted that for the *free access to text with illustrations* condition, readers only have access to one complete version of each illustration (the fourth version of illustration 1 and the fourth version of illustration 2) since it was impossible to predict when readers would choose to inspect an illustration and therefore impossible to coordinate text information with pre sequenced illustration versions.

A series of contrasts were pre-planned to make the following comparisons. First, the free access condition (FA) was compared to the controlled access condition (CA) to determine whether the experimental manipulation of when and where participants were provided with illustration information under the controlled access condition deviated substantially from what would be observed under less controlled (and more natural) reading conditions. Second, conditions with access to both text and illustrations (FA and CA) were compared to the text only condition (TO) to assess the effect of illustrations on text processing. Third, conditions with access to both text and illustrations (FA and CA) were compared to the illustrations only condition (IO) to assess the effect of text on illustration processing. Finally, processing of structural information was compared to the comprehension processing of other types of information (function and energy) to assess the effect of controlling the information perspective from which the domain is described (staging).

Procedure

Presentation of text and illustrations. The text and illustrations were entered into a HyperCard® computer display environment. The environment was designed so that text and illustration information were presented in separate windows. This configuration enabled the experimenter to control if and when text and illustration information would be available to readers and to measure durations of text and illustration consultations. The HyperCard environment was configured to record and time each computer event. This provided a trace of the information that was accessed by each participant as well as a record of the amount of time participants devoted to each information unit (i.e., illustration or text sentence). For each reading condition with text (i.e., TO, FA, and CA), the text was presented to participants one sentence at a time. Participants were able to control the rate of text presentation by pressing a "next" button to obtain the next sentence in the text sequence. Previously read sentences were replaced and did not accumulate on the screen. However, readers were able to look back at previously read text in its entirety by clicking on a button labeled "previous". Similarly, in the illustrations only condition, each illustration was presented separately to participants. Participants were able to control movement between the illustrations by pressing the appropriate button (i.e., "next " or "previous").

The availability of access to the illustration windows varied across conditions. In the text only condition, readers never had access to illustration information. In the free access condition, readers were able to access either of two illustrations from any point in the text.

However, viewing access was limited to one of the two illustrations at any one time. In the controlled access to illustrations condition readers were automatically presented with illustration information at various predetermined points in their reading. There were eight points of access to illustrations for this condition - one for each illustration. Appendix E provides a description where readers in this condition were presented with a particular illustration. Appendix H displays example screen shots of the display environment.

Text and illustration presentation with on-line interpretation. Researchers have established the usefulness of employing concurrent verbalizations as on-line measures of text comprehension and processing. Such measures are particularly useful as indicators of high-level discourse comprehension processes (Lászl6, Meutsch & Viehoff, 1988). For example, Renaud and Frederiksen (1988) have demonstrated that it is possible to investigate on-line local and global semantic processes by measuring sentence reading times, on-line interpretation, and post-input recall. Furthermore, their study demonstrated that varying the on-line interpretation task conditions (i.e., no interpretation, subject-paced interpretation, and experimenter-paced interpretation) did not significantly alter comprehension processing during reading. Deffner (1988) has also demonstrated that the use of a concurrent verbalization task may slow down comprehension processing, but does not qualitatively affect processing. Thus, collecting multiple indicators of text processing, including the use of concurrent verbalization or on-line interpretation tasks, represents an appropriate methodology for research concerned with discourse processing.

The on-line interpretation method was adopted in the present study to investigate how readers use of text and illustrations influenced their comprehension as they read. Readers were able to provide an on-line account of their current understanding of the material at any point in the text and or illustration processing by clicking on an "Interpretation" button. Specifically, participants were instructed to verbalize their thought processes and the content of their current understanding of the human visual system. After completing their response, the reader was returned to the text or illustrations by clicking on the "Next" button. The reader was able to continue reading uninterrupted until the next time they wanted to provide an interpretation. Verbal protocols obtained through this method are assumed to reveal information that is available during comprehension at a given point in text or illustration processing (Ericsson, 1988; Ericsson & Simon, 1980; 1985; Fletcher, 1986; Long & Bourg, 1996; Olson et al., 1984; Suh & Trabasso, 1993; Whitney & Budd, 1996).

Finally, in order to obtain a clear trace of what information readers were focusing on when they consulted an illustration, readers were instructed to use the computer mouse to point to the object(s) of their focus. This information was captured on videotape and subsequently recorded by the experimenter.

Following presentation of the text and/or illustrations, readers were asked to provide both verbal and visual accounts of their understanding of the human visual system based on the materials presented to them. The order of verbal and visual recall was counterbalanced across participants to control for order effects.

Retrospective verbal recall. Participants were asked to orally recall as much information as they could. Specifically, participants were given the following prompt "Tell me in your own words everything that you remember and understood about the human visual system from the materials presented. Provide as much detail as you are able." Retrospective free recalls are assumed to reveal information about what the reader was able to understand and recall, what inferences the reader was able to make, as well as how the reader organized information.

Retrospective visual recall. Readers were also asked to provide drawings of their understanding of the human visual system based on the materials presented to them (i.e., a visual recall). Readers were given the following prompt " Use the pen and paper provided to draw everything that you remembered and understood from all the materials presented to you. You can use the pen to fill in any details or labels of what your remember. Make your drawing(s) as accurate as possible but remember that this task is designed to assess your memory and understanding, and is not a test of your drawing ability." This measure was included to index the participant's ability to understand, recall, and organize information in a visual modality.

Knowledge integration questions. A series of post-reading questions was generated to assess the degree to which readers were able to generate an integrated and coherent understanding of how the visual system functions. Participants were asked to orally answer each question in as much detail as they were able based on their understanding of the material presented to them. These items required readers to make inferential responses to free-response questions. In this sense these items are akin to Mayer's use of a problem-solving transfer task to assess "meaningful learning" (Mayer, 1993a). However, these items also required readers to integrate multiple kinds of text information (i.e., structural, functional, and energy path). A set of scoring procedures detailing the relationships between relevant concepts was constructed for each question. These models were used to assess readers' responses to the questions (see Appendix I for a description of questions)

Dependent Variables

Six different forms of raw data were collected during the experimental sessions. These included: (a) sentence reading times for the groups exposed to text materials (TO, FA, and CA), (b) text look back measures for the groups exposed to the text (TO, FA, and CA), (c) on-line interpretation protocols for all groups, (d) verbal recall protocols for all groups, (e) visual recall protocols for all groups, and (f) answers to post-reading questions for all groups. Each of these forms of data was further manipulated to derive specific measures of interest. The procedures for deriving each of these measures and a description of each measures purpose are described below.

Sentence reading times.[2] Sentence reading times were measured to investigate the information load experienced by participants during the reading task. Each text sentence was coded for the *type* of information it described (i.e., structure, function, energy, or other). By coding the sentences in this way it was possible to investigate whether participants devoted more processing time and resources to one kind of information over another. Appendix E provides a complete list of the categorization of each text sentence according to information type. Reading times for each text sentence were collected and stored automatically by the computer during the reading task. In order to control for possible differences in terms of information density across information types, each sentence reading time was divided by the number of propositions it encoded. Average reading times per proposition for each information type were then calculated for each participant. This yielded three measures of interest that formed the dependent measures subsequently analyzed: average reading times per proposition for sentences describing (a) structure, (b) function, and (c) energy information. These data provide on-line information about the processing load participants experienced for text units describing qualitatively different types of information equated for the overall amount of information encoded.

Text look-backs. Another source of on-line processing data that can be compared across conditions involving text is the frequency, duration, location, and distance of text look-backs made during the reading task. For the groups who were exposed to text (TO, FA, and CA), the number, location, and duration of text look-backs that participants made during the reading task were automatically recorded by the computer. The specific text

[2] Sentence reading times did not include instances of text look-backs.

unit(s) that were consulted on each text look-back were captured on videotape and later recorded by the experimenter for each participant. These look-back behaviors were then coded in terms of (a) the modality of information consulted during the look back (i.e., text only, text and illustrations, illustrations only), and (b) the distances of sentences referred to from the look-back point of origin. Look-backs which involved both text and illustrations were interpreted as indicating instances during which the reader was attempting to integrate information across the two modalities and were coded as instances of *media integration*. Look-backs involving more than one sentence consultation were interpreted as indicating instances during which the reader was attempting to integrate information across several sentences and were coded as instances of *text integration*. Finally, look-backs during which the reader simply went back to the previous sentence just read were interpreted as indicating a specific *comprehension check* or confirmation.

On-line processing measures. On-line protocols were audio taped during the session and later transcribed. The protocols were segmented into main clause segments. Each segment was coded for: (a) content against the text propositions; (b) the type of information expressed (i.e., structure, function, or energy); and (c) the particular on-line processing event that was concurrently demonstrated (see Appendix J for a coded example).

The coding of on-line processing events was based on a modified coding scheme originally developed by Chi et al. (1994) and has been used by other researchers (e.g., Coté, Goldman, & Saul, 1998). The following eight different codes were used: (a) Segments in which the participant provided a paraphrase or literal interpretation of information expressed in the text or illustrations were coded as instances of *interpretation*; (b) Segments in which the participant recalled or referred to prior knowledge and may or may not have directly related this knowledge to the presented content were coded as instances of *prior knowledge* use; (c) Segments in which the participant connected disparate text units were coded as instances of *text integration*; (d) Segments in which the participant connected information presented in the text with information presented in illustrations were coded as instances of *media integration* (this code was limited to participants in the free access and controlled access conditions); (e) Segments in which the participant monitored their comprehension of the material by evaluating the comprehensibility or familiarity of the material, confirmed their understanding, stated a question, expectation or prediction, or identified a comprehension problem were coded as instances of *metacognitive* processing; (f) Segments in which the participant described a particular comprehension strategy they were using or typically use when reading for understanding were coded as instances of *strategy* use. Examples include rereading, visualizing, using prior knowledge to guess or reason about the content, memorizing, skimming, and ordering information; (g) Segments in which the participant commented on characteristics of the text, illustration, or task and were not directly relevant to the content were coded as *comments*; (h) Finally, segments which did not fit any of the above types were coded as *other* (see Appendix K for an example of each code type).

Verbal recall measures. Verbal recalls were audio taped and transcribed. Protocols were segmented into major clause constituents. Protocols were then matched against the text propositions. Each text proposition was also coded for the type of information it encoded (i.e., structure, function, energy, or other). Segments, which matched one or more text propositions, were coded as *recalled* and scored as 1's. In addition, protocol segments which matched text proposition but which also included modifications (addition or substitution) of information were coded as *recalled with inference*. Segments that were coded as recalled with inference were also coded for the type of information they described (i.e., structure, function, energy, or other) according to the original categorization of the proposition. Text segments which did not match any of the text propositions were coded as *inferences*. Inferences were coded for the type of information they described (i.e., structure, function, energy, or other). Also see Appendix L for a coded example.

Thus, it was possible to obtain the following dependent measures to index literal recall: (a) proportion of structural text propositions recalled, (b) proportion of functional text propositions recalled, and (c) proportion of energy propositions recalled. Proportions were calculated to control for differences in the base rate of sentences (and propositions) encoding the different types of information. The same procedure was used to obtain proportions of information recalled with inference and inferred for each information type.

Mental model comprehension. In addition to coding propositional content, each verbal recall protocol was matched to three predefined mental models to index the extent to which participants integrated concepts according to structure, function, and energy organization schemes. Each of these mental models reflects a unique content and structural organization of component parts of the visual system. For example, the *structural mental model* describes a high level spatial arrangement of components and virtually covers all of the component parts of the visual system. Components are connected by "part", "and", and "category" relations. Components that exist within the same general physical space (i.e., outside front, middle, etc.) are grouped together. In contrast, the *functional mental model* describes a high-level process arrangement of components and covers fewer components. Components are arranged according to "part', "temporal order" of processes, "logical conditional", "logical or", and "logical and" relations. Finally, the *energy path mental model* provides a high-level organization of components according to the type of energy they receive and direction of energy flow through the system. Again, fewer component parts are included in this high-level description. These components are connected by "source location", "location order", "resultive location", "conditional", and "goal location" relations. Each of these models provides a unique perspective or conceptual organization of components of the visual system and should therefore be sensitive to measuring the degree to which respondents used each of the perspectives for organizing and integrating their verbal recall of information (see Appendix G).

Visual recall. Visual recalls consisted of one or more drawings made by participants. In order to quantify this data, a list of all the component parts of the human visual system, as described in the text and illustrations, was generated. Each drawing element in participants' drawings was then compared to the list of total elements. Segments which matched one or more of the list of elements were coded as *recalled*. Each drawing element that was coded as present in the protocols was then coded for the *type of information* it conveyed (structure, energy, function, or other).

Drawing elements depicting component parts of the human visual system and their relative location were coded as conveying *structure* information and may have been labeled (score = 1) or not labeled (score = .5) by the participants.

Elements which depicted the source, flow, direction, or type of energy were coded as conveying *energy* information. A score of 1 was allocated for cases in which the energy path information conveyed was accurately depicted and the component part(s) involved was appropriately labeled. A score of .5 was given in cases where the energy information was appropriately depicted but the relevant component part was not explicitly identified.

Annotations to the drawings that described the behavior or function of a depicted component were coded as conveying *function* information. A score of 1 was allocated to cases where the function information was appropriately described and the relevant component(s) involved were appropriately identified. A score of .5 was given in cases describing appropriate functional information but omitting an explicit description of the relevant component(s) involved (see Appendix M for some examples and a listing of the coding categories used).

The total number of drawing elements according to each information type was calculated for each participant. These totals were then converted to proportions based on the total number of possible elements for each information type. Thus, it was possible to obtain the following dependent measures to index visual recall: proportion of (a) structure, (b) function, and (c) energy information recalled

Post-reading questions. Responses to the eight open-ended post-reading questions were transcribed and segmented. Each response was coded against predefined models of the content and structure of the information required to answer each question. None of the questions could be answered based on simply a literal understanding of the stimulus materials. The eight questions were categorized into the following three groups: (a) three questions that required the participant to *integrate information presented in the text only*; (b) three questions that required the *integration of information presented in the text and illustrations;* and (c) two questions that required the *generation of new information*. Scores for each of these three categories of questions were generated by comparing each participant's set of responses to the predefined models. (See Appendix I for the set of questions, scoring models, and coded examples of responses).

Restatement of Experimental Questions and Hypotheses

The following research question and experimental hypotheses were addressed by the study: Each question or hypothesis pertains to particular pre-planned comparisons. Each consists of a set of questions or predictions concerning specific measures of on-line processing, the type of representation generated in comprehending the domain, and the ability to use these representations in responding to questions.

Research question 1. Does providing individuals with free access to illustration information in the context of text (FA) lead to differences in comprehension processing and learning outcomes as compared with the condition with text and experimentally controlled access to illustrations (CA)?

No specific a priori assumptions were made with regard to this question. Several outcomes are possible, however, each of which has different implications. First, it is possible that the two conditions do not differ substantially from one another. This would indicate that the design of the controlled access condition was successful in approximating the manner in which readers naturally use illustration information and did not substantially alter comprehension or learning processes. This in turn would imply that the two text with illustrations conditions could be collapsed for the purposes of other experimental contrasts. Second, it is possible that the CA and FA conditions do differ from one another. Such a finding could have several implications. It is possible that having free access to illustration information affects the efficiency of comprehension and learning but does not substantially affect the quality of final comprehension and learning outcomes. If this is the case, then one would expect to find faster text processing times but equal information recall and mental model comprehension when the FA condition is compared to the CA condition. On the other hand, it is possible that free access affects the quality of overall comprehension. If this were the case, one would expect to find differences in the ability to recall information about the domain, to form mental models of the domain, and/or the ability to answer questions about the domain. Such findings would suggest that readers are more capable of effectively directing their own comprehension processing of materials.

Hypothesis 1. It was expected that when compared to the text only group (TO), the two groups exposed to a combination of text and illustrations (FA and CA) would demonstrate evidence of: (a) more information integration processes in their on-line protocols, (b) the construction of more elaborated and integrated mental model representations of the human visual system, and (c) superior ability to answer comprehension questions that require information integration and the generation of new knowledge in the form of coherent mental models. This hypothesis was based on the assumption that exposure to a combination of text and illustrations facilitates low prior knowledge readers' ability to construct coherent integrated representations (i.e., mental models) of the target domain by providing a constraining visual context in which to understand and reason about text information. No prediction was made concerning the effects on the time devoted to processing text propositions.

Hypothesis 2. It was expected that when compared to the illustrations only group (IO), the two groups exposed to a combination of text and illustrations (FA and CA) would demonstrate evidence of: (a) more on-line elaboration and integrative processing, (b) more complete visual and verbal recalls, (c) more elaborated and integrated mental model representations, and (d) superior ability to answer comprehension questions that require information integration and the generation of new knowledge. This hypothesis follows from the assumption that the text will serve an elaboration function and provide readers with conceptual information necessary to successfully interpret and integrate the information presented through illustrations.

Hypothesis 3. Finally, it was expected that the comprehension processing readers in all experimental groups devote to the *structure* information (i.e., the conceptual perspective from which the materials are described) would be privileged over the processing of the *function* and *energy* information. Specifically, it was expected that, when compared with the function and energy information, participants would demonstrate (a) longer processing times for structure information, (b) proportionally more on-line interpretations describing structure information, (c) greater proportional recall in both verbal and visual modalities of this information, and (d) more elaborated and connected mental models of structure information. Furthermore, it was expected that the effects of information staging would increase the benefits of adding illustrations to text specifically to the comprehension processing of structure information. This hypothesis follows from the staging effects that have been reported in studies of text comprehension and implies that such findings can be successfully extended to the processing of text and illustrations and of illustrations only.

Data Analysis

The design for this experiment is a between-within design. The between-subjects factor is the *presentation* condition with four levels (IO, TO, FA, and CA). The within factor is *information type* with three levels (structure, function, and energy). The data were analyzed using multivariate analysis of variance (MANOVA) models with planned contrasts corresponding to the experimental hypotheses. Repeated measures ANOVAs, MANOVAs, and log-linear analyses were used to investigate each of the main sources of data. Pre-planned contrasts were used to make the following comparisons. First, the free access condition (FA) was compared to the controlled access condition (CA) to determine whether the experimental manipulation of when and where participants were provided with illustration information (under the controlled access condition) deviated substantially from what would be observed under less controlled (and more natural) reading conditions. Second, conditions with access to both text and illustrations (FA and CA) were compared to the text only condition (TO) to assess the effect of illustrations on text processing. Third, conditions with access to both text and illustrations (FA and CA) were compared to the illustrations only condition (IO) to assess the effect of text on illustration processing. Finally, processing of structural information was compared to the comprehension and processing of other types of information (function and energy) to assess the effect of information perspective (staging).

Chapter V: Results

The main research issues addressed in this experiment were: (a) to determine the individual and reciprocal effects of text and illustrations as information sources on comprehension processing, and (b) to determine the effects of controlling the conceptual perspective from which information is described (i.e., staging) on the comprehension processing of text and illustrations. Particular interest was focused on studying the reciprocal effects of media on comprehension processing since past research has not adequately addressed the question of how each source of information is processed in the context of the other. Rather, studies have tended to focus on either the question of how illustrations improve text processing or vise versa. There have been few studies that have attempted to understand how different types of content are processed, and no studies of the effects of information staging on the processing of illustrated text. In the results that follow, emphasis is given to the analysis of conceptual models that participants were able to construct from the presented materials and how such representations are associated with particular patterns of media use.

The following four research questions were addressed: (a) Does providing readers of text with free access to illustration information lead to different comprehension processing and learning outcomes than does controlling the text locations at which the readers are exposed to illustrations?, (b) How do illustrations affect text processing?, (c) How does text affect illustration processing?, and (d) What are the effects of information perspective (i.e., staging) on readers' comprehension processing of text and illustrations? This experiment was designed to provide data relevant to these questions. The study included one between-subjects factor (Information Presentation) with four levels (text only (TO), illustrations only (IO), free access to text and illustrations (FA), and text and illustrations with experimentally controlled access to illustrations (CA) and one crossed within-subjects factor (Information Type) with three levels (structure, function, and energy information).

Several main sources of data were collected including: sentence access times, on-line interpretation protocols, post-input verbal recalls, post-input visual recalls, and answers to a series of open-ended post-input comprehension questions. Each of these data sources provided unique information concerning the content of participants understanding of the domain and the type of comprehension processing they engaged in. For example, on-line reading times provide information about the processing load participants experienced; on-line protocols provide information concerning the content of participants on-going comprehension as well as a trace of the types of processing engaged in; the verbal recall protocols provide information concerning participant's overall understanding of the materials as expressed through verbal descriptions, the inferences they made, and the overall way in which they structured that information; the visual recall protocols provided an alternate format for participants to express their overall understanding, and finally, the post-input questions provided data about how participants were able to integrate certain concepts that were described in the text or through a combination of text and illustrations as well as their ability to generate new knowledge.

A series of MANOVAs, repeated measures ANOVAs, and log-linear analyses were used to investigate each of the main sources of data. Pre-planned contrasts were used to make the following comparisons:

(a) The free access condition (FA) was compared to the controlled access condition (CA) to determine whether the experimental manipulation of when and where participants were provided with illustration information under the controlled access condition deviated substantially from what would be observed under less controlled (and more natural) reading conditions.

(b) Conditions with access to both text and illustrations (FA and CA) were pooled and compared to the text only condition (TO) to assess the effect of illustrations on text processing.

(c) Conditions with access to both text and illustrations (FA and CA) were pooled and compared to the illustrations only condition (IO) to assess the effect of text on illustration processing.

(d) Each measure of the processing of *structural* information was compared to the average of the other two types of information (*function* and *energy*) to assess the effect of information staging.

Generally speaking, a main effect of *presentation condition* would indicate that source of information leads to significant differences in comprehension processing. A main effect for the planned contrast involving *information type* would indicate that comprehension processing differs depending on the staging of the information perspective presented. Finally, an interaction of *presentation condition* and *information type* would provide evidence that differences in comprehension processing of various information types depend on the modality of information presentation.

The reporting of results will be organized under two main themes: (a) the analysis of *on-line measures* of comprehension processing, and (b) the analysis of *post-input measures of comprehension processing*. In general, the on-line measures provide information about the manner in which participant's processed the learning materials, and are sensitive to minute variations in comprehension processing that are closely linked temporally to the participant's behavior. The post-input measures provide information about each participant's overall comprehension, organization, and integration of their conceptual representations of the target domain.

On-line Processing of Text and Illustrations

Processing times. The first research question addressed concerns those participants who were exposed to text information (i.e., TO, FA, and CA). The specific issues raised are (a) whether the presence of illustrations affects the degree of processing effort (i.e., time) that participants exert in trying to comprehend text information, (b) whether information staging affects processing time, and (c) whether or not the effects of staging on processing time varies according to whether or not illustration information was available. It was expected that readers would devote more processing resources to text sentences describing information from the perspective that was staged high (i.e., structure information) than to other information (i.e., function and energy information).

Sentence reading times were analyzed to investigate the information processing load experienced by participants during the reading task. As described in the methods section, average reading times per proposition for each information type were calculated for each participant. This yielded three measures of interest that formed the dependent measures subsequently analyzed: average reading times per proposition for sentences describing (a) structural, (b) functional, and (c) energy information. These data provide on-line information about the processing load participants experienced for text units describing qualitatively different types of information equated for the overall amount of information encoded.

The reading time data were analyzed using a 3 X 3 repeated measures MANOVA consisting of one between subjects factor (presentation condition) and one within subjects factor (information type). The MANOVA test results revealed that there was a significant main effect of information type (multivariate $F(2, 14) = 49.511, p < .0005$). As expected, readers spent proportionally more time reading the structural information than they did reading functional or energy path information ($F(1, 15) = 86.130, p < .0005$). That is, participants exerted more processing effort when reading information that was highlighted as more important in the text structure. A contrast for the condition effect was performed to compare the performance of the free access condition to the controlled access condition. This comparison did not yield a significant difference between the two groups who were exposed to both text and illustrations ($F(1,10) = 1.164$, $p = .306$). However, there was a significant main effect of condition for the contrast between the combined reading conditions with access to illustrations (FA and CA) and the text only group ($F(1,15) = 16.4$, $p = .001$). The groups who were provided with illustrations read text sentences significantly faster than subjects who were only exposed to the text alone (see Figure 3 below). Finally, this effect did not vary across the different types of information (multivariate $F(4, 28) = 0.666$, $p = .621$).

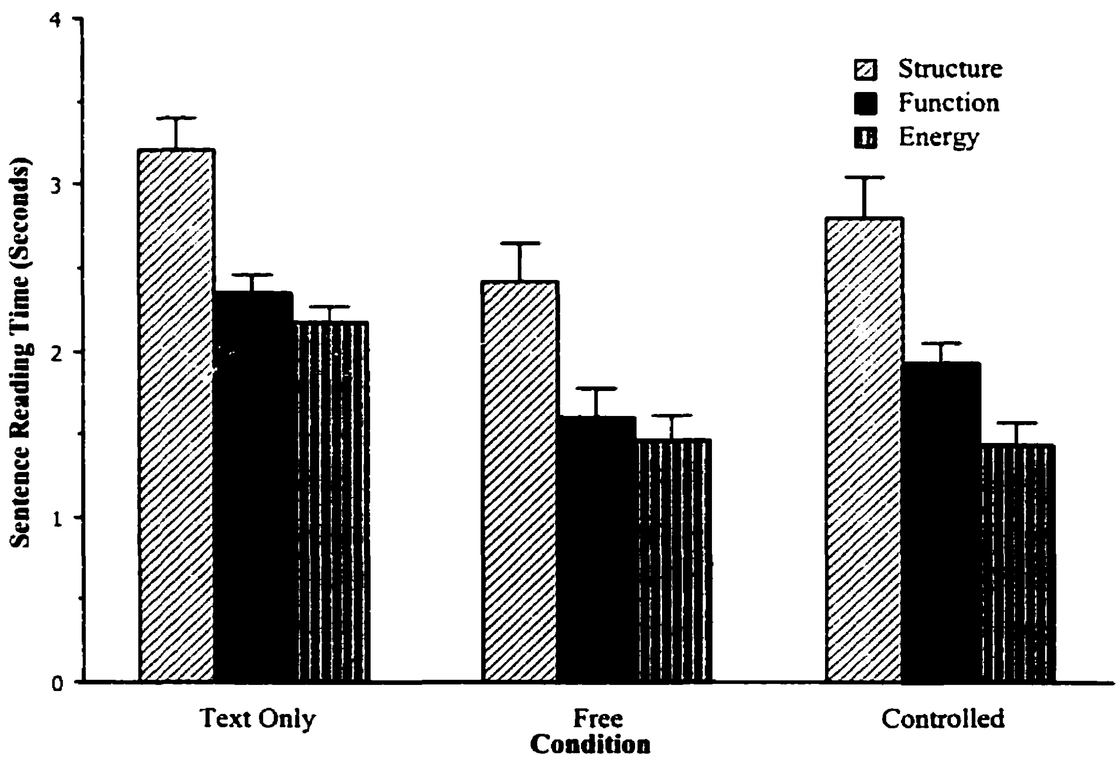


Figure 3. Average sentence reading times by condition and information type controlled for number of embedded propositions.

Why was it that the groups exposed to a combination of text and illustrations processed text information faster than the text only group? To investigate the question further, a new dependent measure was created indexing the *total amount of time* participants spent processing the experimental materials. For the text only group this measure consisted of the total reading time each participant devoted to reading all the text sentences. For the FA and CA groups this measure consisted of the total time devoted to reading the illustrations plus the total amount of time devoted to reading all the text sentences. A one-way ANOVA with contrasts revealed that the two text with illustrations groups did not differ from one another in terms of total processing time ($F_{(1,15)} = 1.601$, $p = .234$). Furthermore, the TO group did not differ significantly from the combined text with illustrations groups (FA and CA) on this measure ($F_{(1,15)} = 0.462$, $p = .509$). The means for the total processing time by condition are presented in Figure 4.

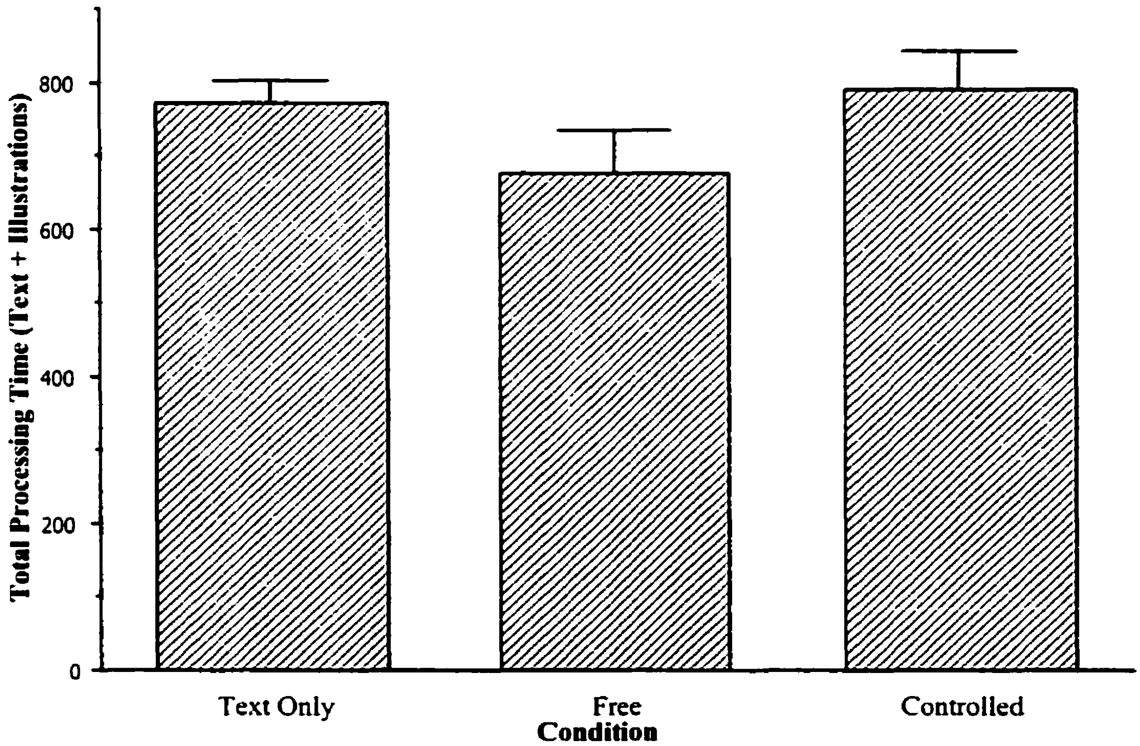


Figure 4. Total processing time for text and illustration information by condition.

Text look-backs. To explore possible group differences in terms of the pattern of text look-backs during the reading task the frequency, duration, and distance of text regressions were analyzed. The frequency and location of each text look-back was automatically recorded by the computer. Each text look-back was later categorized by the type of information from which the look-back was initiated by the reader. A repeated measures analysis was conducted to investigate differences across conditions, differences across information types, and an interaction of these two variables. This analysis revealed a significant main effect for information type ($_{\text{multivariate}}F_{(2,14)} = 10.261$, $p = .002$).

The mean number of text regressions for *structural* information was greater than the average number for function and energy information ($F(1,15) = 20.47, p < .0005$). Figure 5 below presents the average frequencies of text regressions by condition by information type.

The duration of each text look-back was automatically recorded by the computer. A one-way ANOVA was conducted on the average time each participant spent per text look-back across conditions. This analysis revealed that the mean look-back duration of the text only group was greater than the mean of the two text with illustrations groups on the average duration per text look-back measure ($F(1,15) = 11.21, p = .004$). The two text with illustrations groups did not differ from one another on this measure ($F(1,15) = 0.002$, $p = .969$). Figure 6 displays the average time spent on look-backs by condition.

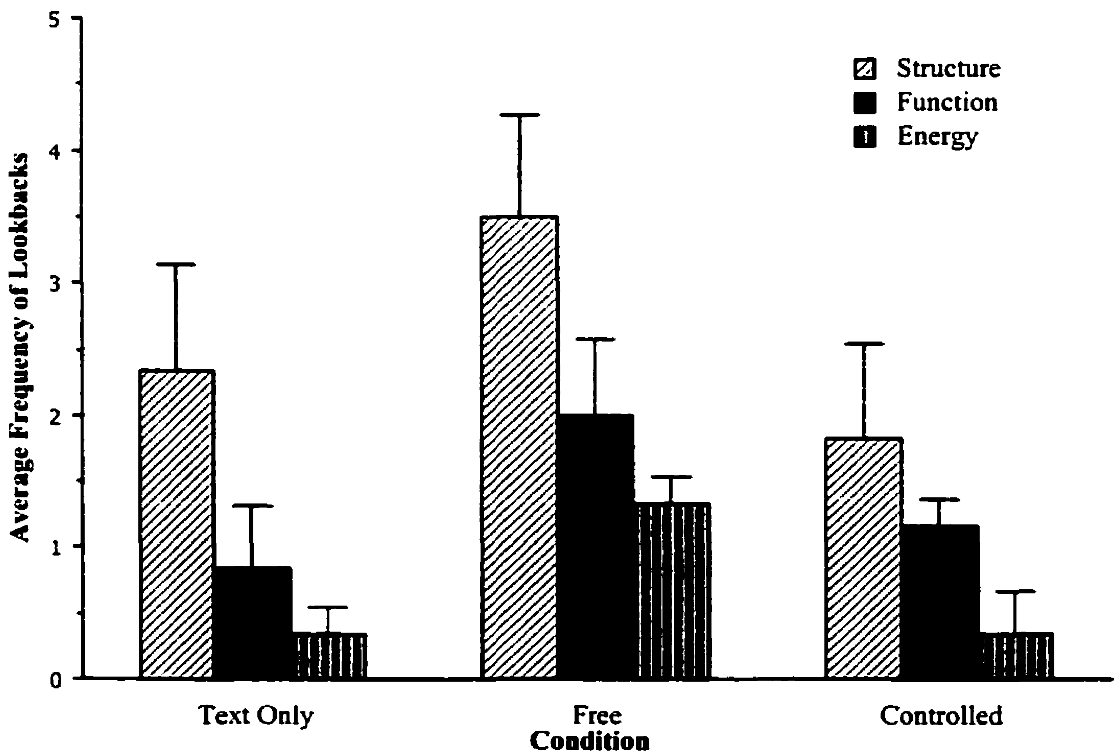


Figure 5. Average frequency of text look-backs by condition by information type.

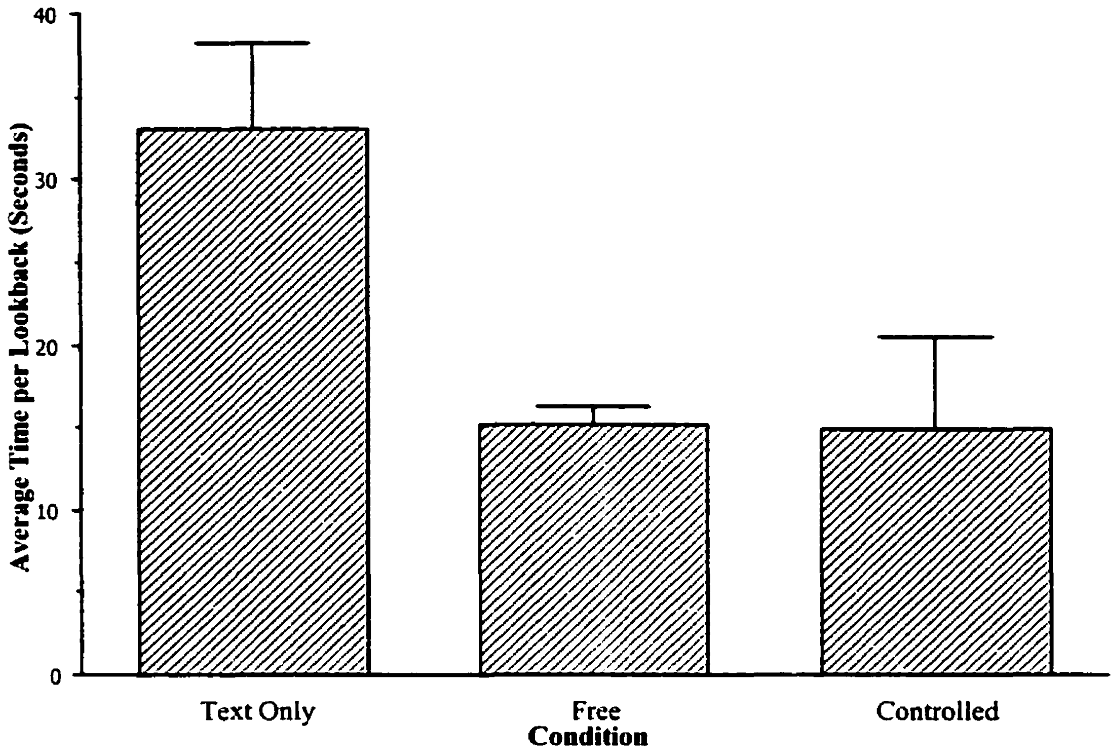


Figure 6. Average duration per text look-back by condition.

The average distance of look-backs was calculated for each subject as described in the methods section. When the free access and controlled access conditions were compared on this measure they did not differ significantly ($F(1,10) = .337, p = .574$). Together, however, the two pooled text with illustrations groups did differ significantly from the text only group ($F(1,16) = 4.556, p = .049$), with the text only group demonstrating significantly further look-backs. Thus, the text only group spent significantly longer and traversed greater text distances than the text with illustrations groups. Appendix M also presents a descriptive summary of the look-backs coded for type (i.e., comprehension checks, text integration, and media integration) by condition.

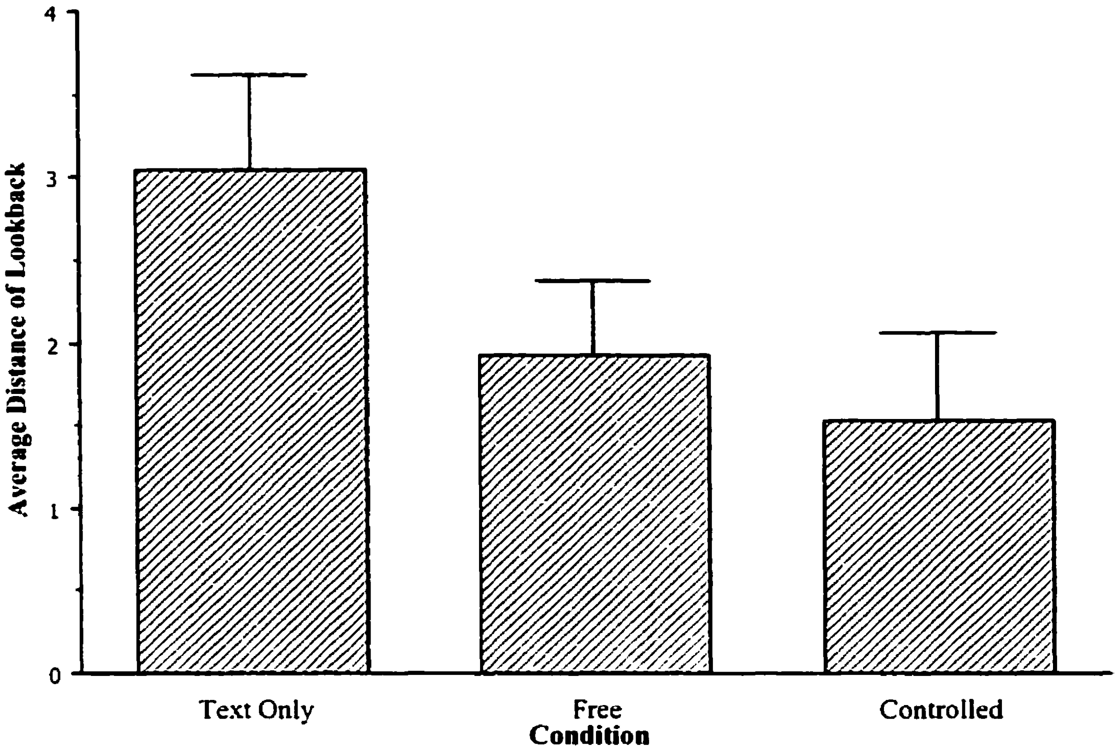


Figure 7. Average distance per text look-back by condition.

On-line content. In order to investigate readers' on-line processing of information content, each participant's on-line protocol was coded for the type of information expressed (i.e., structure, function, energy, or other). Based on this coding it was possible to derive the proportion of each participant's protocol that expressed (a) structure, (b) function, and (c) energy descriptions. These measures were analyzed to investigate the effects of presentation conditions and information staging on participants' on-line processing of the propositional content.

A repeated Measures MANOVA revealed a main effect of information type for the groups that were exposed to text in which reader's uniformly made proportionally more on-line statements describing structural information than functional or energy information ($F(1, 20) = 10.429$ $p = .004$). There were no differences between the text with illustrations groups and the text only group on these measures. However, the illustrations only group demonstrated a different pattern of on-line processing. Participants in this condition focused on describing both structure and energy information. The contrast between the illustrations only condition and pooled text with illustrations conditions yielded a significant difference for the processing of energy information ($F(1,20) = 4.916$,

$p = .038$). Specifically, the IO group demonstrated significantly more on-line interpretations describing energy information than the groups that were exposed to texts with illustrations. This was an unanticipated result given that an attempt was made to privilege structure information over energy information in the illustrations as well as in the text.

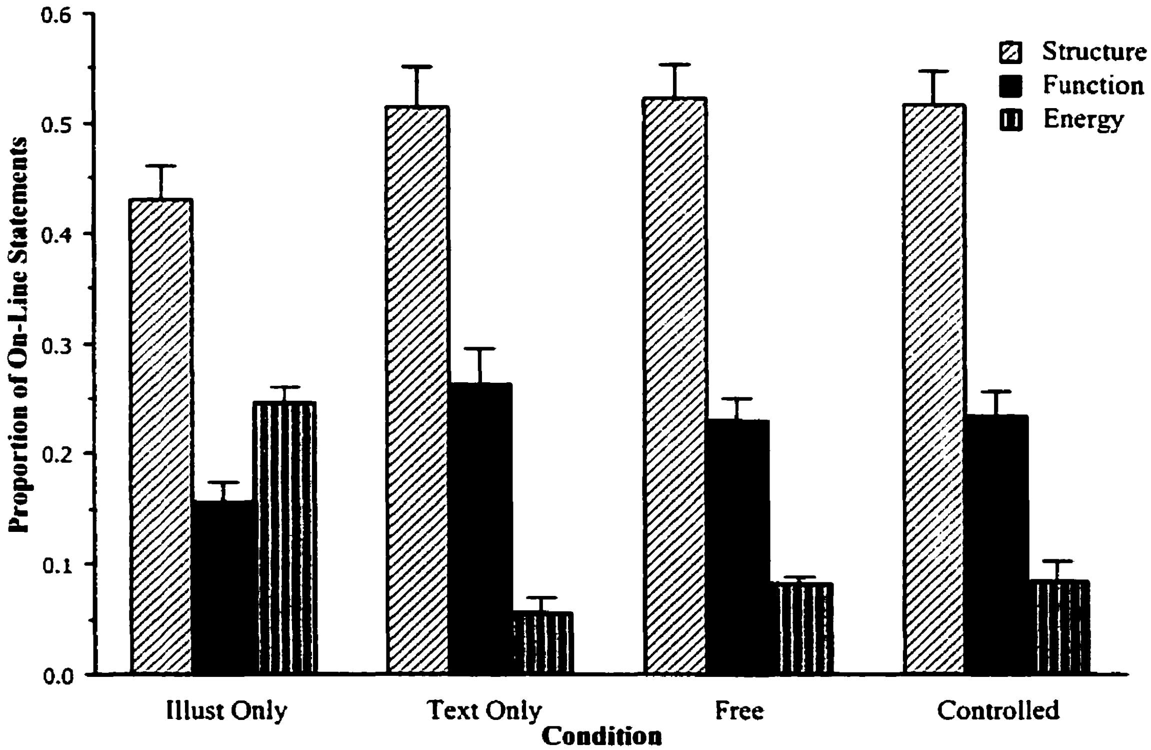


Figure 8. Proportion of on-line statements by condition and information type.

On-line interpretation events. As described in the methods section, each on-line processing event that occurred during the presentation of the learning materials was coded and categorized for the particular type of comprehension processing demonstrated (e.g., interpretation, inference, strategy, etc.). Total frequencies for each processing category type were then computed for each presentation condition. Two separate log-linear analyses were performed to examine whether the profile of frequencies of on-line processing events differed according to presentation conditions. The first log-linear analysis was a 2 X 6 and included condition (the pooled text and illustration groups (FA and CA) vs. the text only group) crossed with the six different on-line processing codes (*comment*, *inference*, *integration*, *interpretation*, *metacognition*, and *strategy*). The results revealed a significant interaction between type of on-line event and condition (χ^2 (5) = 11.355, p =.045) demonstrating that the profile of on-line comprehension processing was significantly different across the two conditions. Table 9 below describes the percent of each type of processing event observed within each condition. As can be seen in the table, the text and illustrations groups demonstrated proportionally more inferences, integration, metacognition, and strategy use than the text only group. Thus, even though these two groups were reading the text information significantly faster than the text only group, they were also demonstrating a more varied pattern of comprehension processing.

Table 9

Row Percent of On-line Processing Events by Condition (FA & CA vs. TO)

On-line Processing Event

Condition	Comment	Inference	Integration	Interp	Metacog	Strategy	Total %	N
FA & CA	8.33	14.10	14.74	21.15	16.67	25.00	100	156
TO	14.29	7.14	9.52	42.86	11.90	14.29	100	42
Total %	9.60	12.63	13.64	25.76	15.66	22.73	100	
N	19	25	27	51	31	45		198

Note. FA & CA = pooled text with illustrations conditions; TO = text only condition; Interp = interpretation; Metacog = metacognition.

A second 2 X 6 log-linear analysis was performed to compare the profile of on-line processing events observed for the text with illustration groups to the illustrations only group. The analysis revealed a significant interaction between category of on-line events and condition (χ^2 (5) = 157.209, $p < .0005$). Table 10 below presents the percent of each type of processing event observed within each condition. Again, the text with illustrations conditions demonstrate heavier use of inferencing, integration, metacognitive processing, and strategy use than the illustration only condition during the presentation of learning materials. Thus, qualitatively distinct patterns of on-line processing were associated with the text with illustration groups as compared to the text only group, and as compared to the illustration only group.

Table 10

Row Percent of On-line Processing Events by Condition (FA & CA vs. IO)

On-line Processing Event

Condition	Comment	Inference	Integration	Interp	Metacog	Strategy	Total %	N
FA & CA	8.33	14.10	14.74	21.15	16.67	25.00	100	156
IO	7.58	0.00	6.11	68.72	7.58	0.00	100	211
Total %	7.90	5.99	15.53	48.50	11.44	10.63	100	
N	29	22	57	178	42	39		367

Note. FA & CA = pooled text with illustrations conditions; IO = illustrations only condition; Interp = interpretation; Metacog = metacognition.

In order to investigate whether this different pattern of on-line processing between conditions was affected by information type a log linear analysis was conducted which included three factors: Condition, Processing Event, and Information Type. In an attempt to overcome the problem of sparse expected frequencies the coding categories were collapsed in the following manner. First for the dimension of information type, two categories were created. Category one included all instances of on-line interpretations describing structure, and category two included pooled instances of on-line interpretations describing function and energy information. For the dimension of on-line interpretation events two categories were created. Category one included all instances of *"active" on-line processing* (i.e., pooling inferences, integration, metacognition, and strategy categories). Category two included all other types of coded on-line events (i.e., pooling the comments and interpretation categories). Even with collapsing the data in this way, however, it was still not possible to generate adequately large expected frequencies to obtain stable estimates of effects across the different information types.

When the frequency counts were restricted to interpretations which described *structural* information, however, the expected frequencies proved to be high enough to produce stable estimates when the processing events were collapsed into three categories: comments, interpretations, and "active" processes (which consisted of a collapse of inferences, integration, metacognition, and strategy use). In this way, it was possible to perform a third two-way log-linear analysis on the on-line processing of *structural* information alone to investigate whether distinct processing profiles of on-line comprehension processing events were associated with different conditions for the type of information that was staged high. This analysis revealed a significant interaction (χ^2 (4) = 65.881, $p < .0005$). Table 11 below describes the distribution of processing events within each condition. The table clearly indicates that conditions with text and illustrations were associated with distinct (specifically *more active*) patterns of on-line processing of structural information than either the text only or illustration only conditions.

Table 11

Row Percent of On-line Processing Events for Structural Information by Condition

On-line Processing Event

Condition	Commen	Interpretatio	Active	Total %	N
Text &	5.68	29.55	64.77	100	88
Text Only	9.09	72.73	18.18	100	22
Illustrations Only	0.00	83.51	16.49	100	97
Total %	3.38	59.42	37.20	100	
N	7	123	77		207

Note. Active = pooled instances of inferences, integration, metacognition, and strategy categories.

Summary on-line results. To summarize, there were three main results obtained for the on-line data. First, the two groups who were exposed to a combination of text and illustrations did not differ from one another in their on-line processing of information. Second, the pattern of on-line comprehension processing for the two pooled text with illustrations groups did differ from the text only group. Specifically, the two groups who were exposed to text with illustrations processed text information much faster, with shorter text regressions, but were more active in their on-line interpretation of information than the text only group. The three groups that were exposed to text, however, did *not* differ from one another in terms of total processing time, only in how that time was allocated. Each of the groups that were exposed to text demonstrated a classic staging effect in their on-line processing of the materials. That is, each of these groups devoted more time, made proportionally more on-line interpretations, and made more frequent text look-backs in processing information that was staged high (i.e., structural information). Third, the groups that were exposed to text with illustrations differed from the group that was exposed only to illustrations in terms of the effect of information staging. While the only on-line data that was available for the illustrations only group was their on-line protocols, these data demonstrated that these individuals focused on both structural and energy information. In addition, these data indicated that individuals in the IO condition were less active in their on-line processing of information than were those in the text with illustrations groups.

Post-Input Measures of Text and Illustration Comprehension

While significant differences were observed among the presentation conditions in the on-line processing of the text and illustrations, the consequences such differences had on participant's overall understanding of the domain also needs to be investigated. To address this issue we turn now to examine the results of analyzing the post-input data.

Propositional recall. As described in the methods section, a propositional analysis was performed on the text to provide a representation of the semantic information that was literally encoded in the text. Each text proposition was then categorized for the type of information it encoded (structure, function, energy, other). This analysis provided a model against which each participant's verbal recall could be matched. Based on this analysis it was possible to derive the following three measures for each participant: (a) proportion of structural text propositions recalled, (b) proportion of functional text propositions recalled, and (c) proportion energy text propositions recalled. These measures formed the dependent variables used in a repeated measures MANOVA to assess the effects of presentation condition and information type on participants' ability to literally comprehend and recall propositional information presented in the text.

There was no overall significant difference in terms of literal recall of propositional information for the contrast between the free access (FA) and controlled access to illustrations (CA) conditions ($F_{(1, 20)}$ =,001, p = .973). In addition, the text only condition did not differ from the pooled text with illustrations conditions on this measure ($F_{(1, 20)}$ = 0.669 p = .423). Furthermore, there was no preference given to recall of structural information over and above recall for function and energy information for these groups ($F_{(1,20)}$ =0.000, p = .993). However, when the illustrations only condition was contrasted with the pooled text with illustrations conditions there was a significant difference as anticipated ($F(1,20) = 7.243, p = .014$). Interestingly, further univariate analysis revealed that these differences were specific to recall of structural ($F(1,20) = 5.434, p =.030$) and functional information ($F(1,20) = 14.402, p = .001$). Curiously, however, the two groups did *not* differ in their ability to recall energy information ($F(1,20) = .176, p = .680$). Figure 9, below displays the average percent of information recalled verbally by condition and by information type.

In sum, there were no differences between any of the presentation conditions that were exposed to text in their ability to provide literal recall of text propositions, nor were there any effects involving information staging for these conditions. On the other hand, the illustrations only group did demonstrate significantly poorer recall of both structural and functional propositions when compared to the text with illustrations groups. However, they did not differ significantly in terms of their ability to recall energy information, even though they had not been exposed to these propositions in the text.

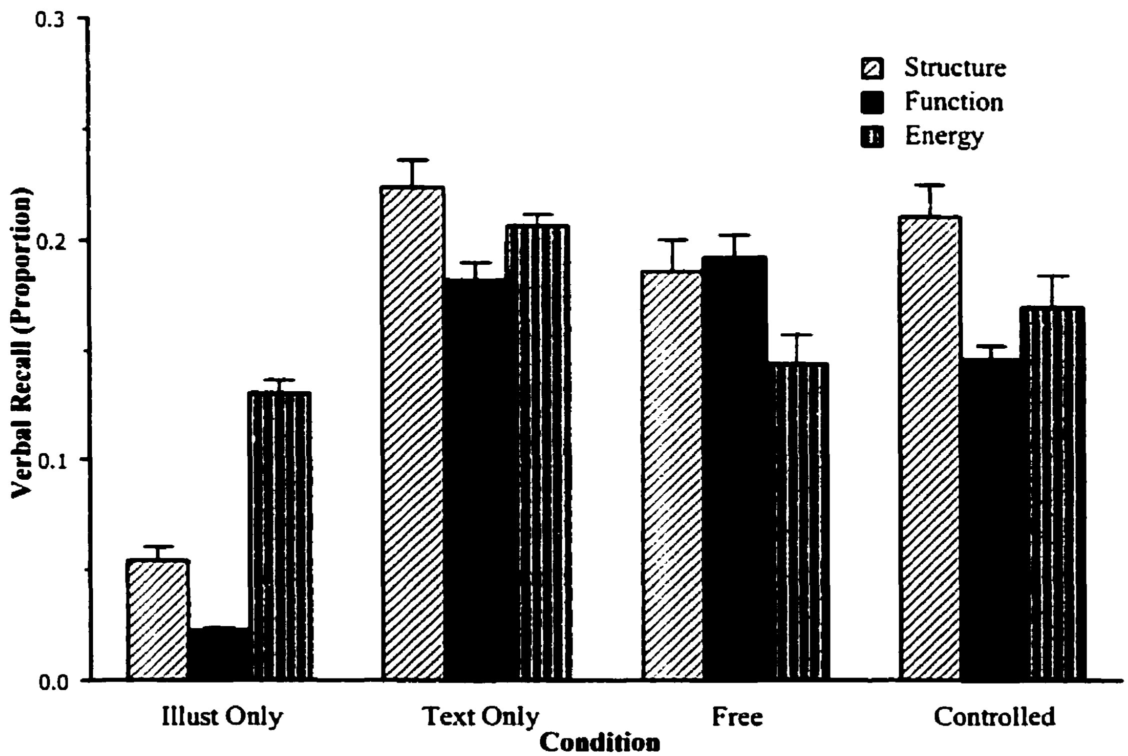


Figure 9. Average proportion of text propositions verbally recalled by condition and information type.

Staging of topics in participants' verbal recall protocols. In order to investigate the IO group's behavior more closely, the opening segments of each protocol were examined. This examination revealed that the majority of participants in the illustrations only group began their recalls with energy path descriptions. Essentially, these individuals introduced descriptions of components of the human visual system in the context of *energy path information*. Five of the six participants in the IO group demonstrated this pattern. In contrast, individuals in each of the conditions involving text mostly opened up their verbal recalls by introducing and organizing information from the point of view of *structural information*. Although, this analysis is purely descriptive in nature, it does suggest that participants in the IO condition were using energy information to provide a narrative structure for organizing their description of structure and function information, and staged energy information high in their discourse. Appendix M provides a typical example of verbal recall openings for each condition.

Structural inferences. Protocol segments which did not match text propositions were coded as inferences. Each inference was coded for the type of information it described (structure, function, or energy). The total number of inferences for each information type was then converted to a proportion of each individual's total number of protocol segments. There were no differences between the FA and CA groups on this measure ($F(1,20) = 0.756, p = .395$), nor was there a difference between the two pooled FA and CA groups and the TO group ($F(1,20) = .120, p = .733$). The two pooled FA and CA conditions, however, did generate significantly more inferences than the IO group ($F(1,20) = 11.391, p = .003$). Further analysis revealed that the two pooled FA and CA groups made significantly more inferences concerned with *structural* ($F(1,20) = 5.309, p = .008$) and *function* information ($F(1,20) = 10.623, p = .004$) than the IO group, but did not differ on the proportion of inferences generated for *energy* information ($F(1,20) = 2.360, p = .140$). See Figure 10 .

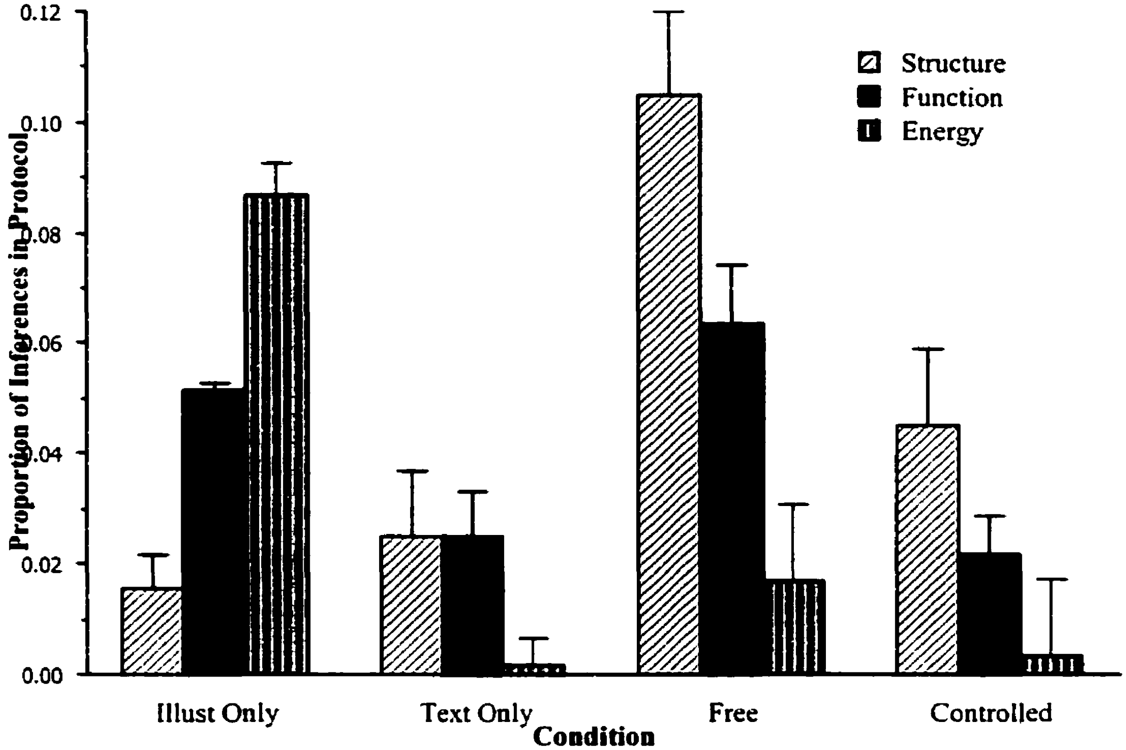


Figure 10. Proportion of inferences in verbal recalls by condition and information type.

An exploratory analysis, however, did reveal significant differences in terms of the types of structural inferences generated. Two types of structural inferences were distinguished. First, inferences which involved descriptions concerned with the relative location of components were coded as *location* inferences. Second, inferences which described the number, density, part, or connection among components were coded as *arrangement* inferences. Table 12 below provides an example of each type of inference.

Table 12
Example of Location and Arrangement Inferences Describing Structural Information

Condition	Protocol Segment	Information Type	Process	Inference Type
CA	Umm there's the pupil,	structure	recall	X
	which is the opening in the iris,	structure	recall	X
	which lies behind the cornea,	structure	inference	location
	which is in front of the lens.	structure	inference	location
TO	umm the the photoreceptors are not evenly distributed.	structure	inference	arrangement
	There's a region of highest concentration called the fovea.	structure	inference	arrangement

Note. The example of a location inference is taken from a participant in the text with controlled access to illustrations (CA) condition. The example of an arrangement inference is taken from a participant in the text only (TO) condition. X = no inference.

The proportion of segments in each individual's protocol describing arrangement and location type inferences was calculated. A one-way MANOVA of the proportion of these types of structural inferences by condition revealed that the two text with illustrations groups did not differ from one another (multivariate $F(2, 19) = 0.185, p = .832$). However, the pooled text with illustrations conditions displayed a significantly different pattern of structural inferences than the text only group (multivariate $F(2, 19) = 5.341, p = .014$). Further univariate analysis revealed that the text only group made significantly more *arrangement* inferences ($F(1, 20) = 7.674, p = .012$) while the two text with illustrations groups made more *location* inferences ($F(1, 20) = 4.549, p = .046$). Interestingly, many of the location inferences that were made by the text only group were erroneous (e.g., "And on top of the cornea is the sclera."). As Figure 11 below indicates, the illustrations only group did not make many structural inferences at all. The text with illustrations groups made significantly more structural inferences in general than the illustrations only group (multivariate $F(2, 19) = 6.110, p = .009$). Univariate tests revealed significant differences on both arrangement ($F(1, 20) = 6.464, p = .019$) and location inferences ($F(1, 20) = 5.238, p = .033$).

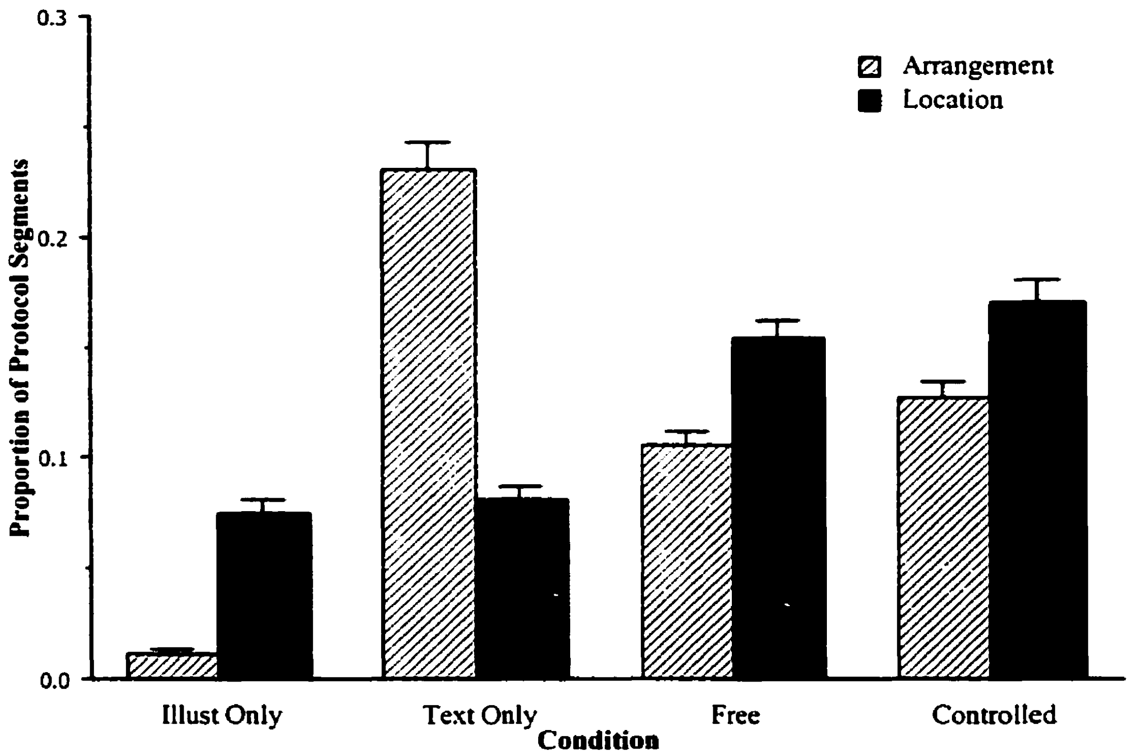


Figure 11. Average proportion of structural inferences in verbal recalls by condition and inference type.

Mental model comprehension. In addition to coding propositional content and inferences, each verbal recall protocol was matched to three predefined mental models to index the extent to which participants integrated concepts according to structure, function, and energy path organization schemes. In order to assess the completeness of participants' high level comprehension, the proportion of *mental model nodes* covered in each verbal recall protocol was calculated for each type of mental model (i.e., structure, function, and energy). Analysis of these data revealed that the two text with illustrations groups (CA and FA) did not differ from one another (multivariate $F(3,18) = 0.348, p = .791$). The contrast between the text only group and the pooled text with illustrations conditions revealed a significant difference (multivariate $F(3,18) = 3.514$, $p = .036$). As anticipated, the text with illustrations groups were more successful than the text only group at constructing mental models of information from a structural perspective ($F(1, 20) = 9.274$, $p = .006$). The groups did not differ in terms of their ability to form mental models of either function ($F(1,20) = 0.843$, $p = .369$) or energy information ($F(1,20) = 0.098$, $p = .757$). As anticipated the illustrations only group differed from the pooled text with illustrations groups (multivariate $F(3,18) = 7.139$, $p = .002$). The illustration only group differed significantly from the text with illustrations groups in terms of their ability to elaborate both structural ($F(1,20) = 13.514$, $p = .001$) and functional mental models ($F(1,20) = 10.599$, $p = .004$), but not in their elaboration of energy mental models ($F(1,20) = 2.138$, $p = .159$).

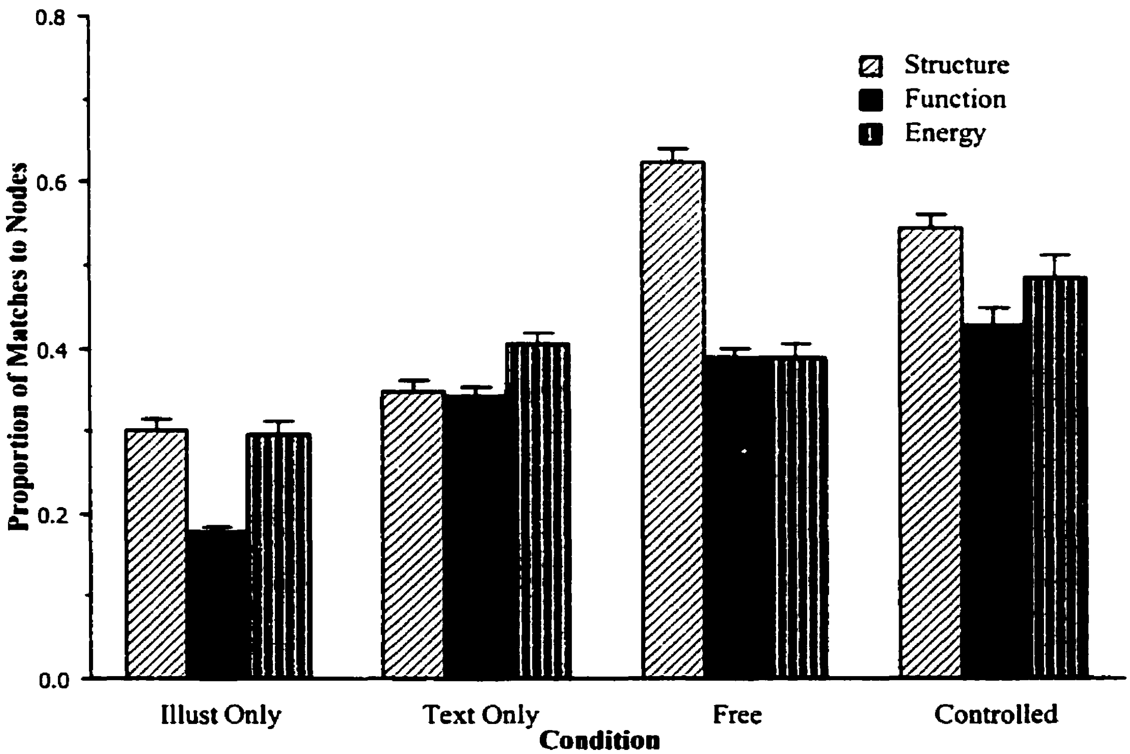


Figure 12. Average proportion of elements matched to mental models by condition and information type.

In addition to the proportion of mental model nodes matched, participants' protocols were also scored in terms of the proportion of *mental model relations* described for each information type. This measure was included to index the degree to which model elements were explicitly connected by each participant. This analysis yielded a similar pattern of results as above. The two text with illustrations groups did not differ from one another (multivariate $F(3, 18) = 0.390, p = .762$). The text with illustrations groups differed than the text only group in terms of the degree to which they developed connected mental models (multivariate $F(3,18) = 6.649, p = .003$). Further univariate analysis revealed that the text with illustrations groups described more connected mental models of structural information than did the text only group ($F(1,20) = 17.628, p < .0005$). The text with illustrations groups also differed significantly from the illustrations only group on this set of measures (multivariate $F(3, 18) = 12.618, p < .0005$), demonstrating more integrated mental models for structural ($F(1, 20) = 31.340$, p < .0005), functional ($F(1, 20) = 5.336$, p = .032), and energy perspectives ($F(1,20) = 7.827, p = .011$).

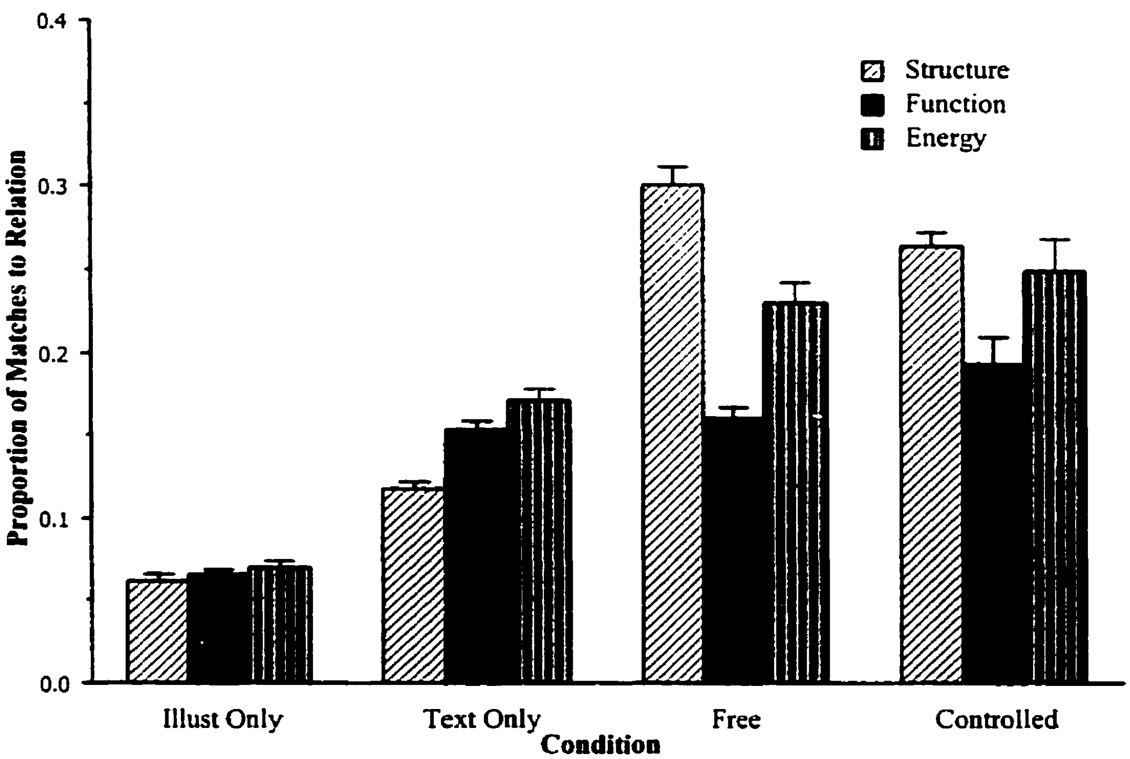


Figure 13. Average proportion of relations matched to mental models by condition and information type.

Mental model comprehension controlling for processing effects. The question of why the combination of text and illustrations outperformed the text only group in terms of their ability to build and elaborate mental models of structural information but not in terms of their literal recall needs to be addressed. One hypothesis that is raised by the pattern of results so far is that mental model performance is positively related to the degree of active on-line interpretation. To investigate this possibility, two separate MANCOVAs were used to assess the effect of presentation condition on the proportion of (a) mental model nodes and (b) mental model relations covered while controlling for differences on the proportion of active on-line processing events (that were previously established as differentiating the groups).

The difference between the text only (TO) and the pooled text with illustrations conditions (FA and CA) on the proportion of mental model *nodes* covered was attenuated (multivariate $F(3,17) = 1.663, p = .213$) when the proportion of active on-line statements was used as a covariate. Univariate analysis also indicated that the previously significant difference between these groups in terms of mental model elaboration of structural information was also attenuated ($F(1,19) = 2.244, p = .151$). In addition, the difference between the illustrations only group (IO) and the pooled text with illustrations groups (FA and CA) was also non significant when active on-line interpretations were controlled for (multivariate $F(3,17) = 2.801, p = .071$). None of the univariate differences were significant.

The second MANCOVA indicated that the difference due to the contrast between the TO and the pooled CA and FA conditions on the proportion of mental model *relations* covered (multivariate $F(3,17) = 4.636$, $p = .015$) and the contrast between the IO and the pooled CA and FA conditions (multivariate $F(3,17) = 7.301$, $p = .002$) were only partially attenuated when the degree of active on-line interpretations was controlled. Both effects remained significant, although they did decrease.

Tests of the assumption of parallelism of regression for both MANCOVAs were met, indicating that nature of the relationship between the active on-line processing variable and the mental model node and mental model relation variables was consistent across conditions.

In addition, simple correlations and partial correlations controlling for the effects of presentation condition were run between the degree of "active" on-line interpretations made during the reading task and the proportion of mental model nodes and relations elaborated in participants' protocols. The simple correlation analysis revealed a significant correlation between the proportion of each participant's use of active on-line comprehension processing and the proportion of matches to both nodes and relations for structural information at the mental model level (see Table 13 below). This relationship did not carry over for function or energy information. When the group differences due to presentation condition were partialled out, no significant relationships remained between the active on-line interpretation variable and the mental model nodes and relations variables. Therefore, the effects of the processing variable were due to *condition* effects on the processing variable, and not to individual variation in active processing.

Table 13

Correlations Between Proportion of "Active" On-line Interpretations and Proportion of Mental Model Nodes and Relations in Participant's Verbal Protocols by Information Type

	Proportion of Structure Nodes	Proportion of Function Nodes	Proportion of Energy Nodes
Proportion of active on-line interpretations	0.597**	0.392	-0.020
Proportion of active on-line interpretations corrected for condition effects	0.249	0.081	-0.172
	Proportion of Structure Relations	**Proportion of Function Relations**	**Proportion of Energy Relations**
Proportion of active on-line interpretations	0.690**	0.311	-0.169
Proportion of active on-line interpretations corrected for condition effects	0.197	0.089	-0.184

Note * $\alpha < .05$; ** = $\alpha < .001$

Visual recall. Each component part of the visual system that was described in participants' visual recall protocols was coded for the type of information described.

Proportion of information recalled for each participant was then calculated for each information type. Analysis of these data revealed no significant difference between the two text with illustrations groups (multivariate $F(3,18) = 2.106, p = .135$). There also was no significant main effect for the contrast between the text only group and the groups exposed to a combination of text with illustrations ($F(3,18) = 2.356, p = .106$).There was a significant main effect of information type for all of the groups exposed to the text (multivariate $F(2,19) = 62.503, p < .0005$) whereby recall for structural information was significantly higher than for function and energy information ($F(1, 20) = 75.348, p < .0005$). There was, however, a significant interaction ($F(1, 20) = 6.724, p = .017$) between information type (structure vs. function and energy) and condition (two pooled text with illustrations groups vs the illustrations only group). As indicated in Figure 14, the illustrations only group recalled more energy information (than sturcure or function information) whereas the two text with illustrations groups recalled more stucture (than energy or function information). Thus, these results provide further evidence that the IO group demonstrated preference for energy path information.

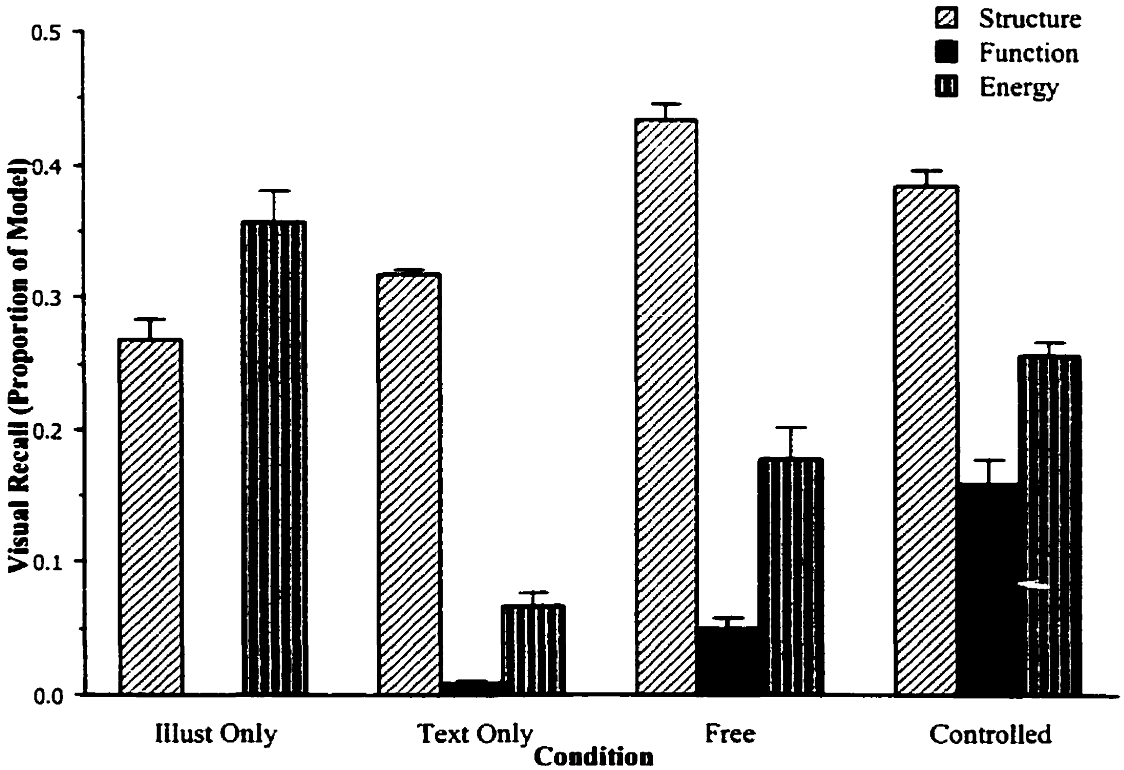


Figure 14. Average proportion of information recalled visually by condition and information type.

Visual recall errors. In addition to the visual recall data, the proportion of errors made in the visual recall protocols was recorded. These errors included instances of (a) location errors, (b) logical errors, and (c) label errors. The effect of presentation condition on the total proportion of errors made was investigated using a one-way ANOVA. The text only group made significantly more errors than the text with illustrations group ($F(1,20) = 10.748$, $p = .004$). In addition, the illustrations only group made proportionally more errors than the text with illustrations group ($F(1,20) = 4.795$, $p = .041$).

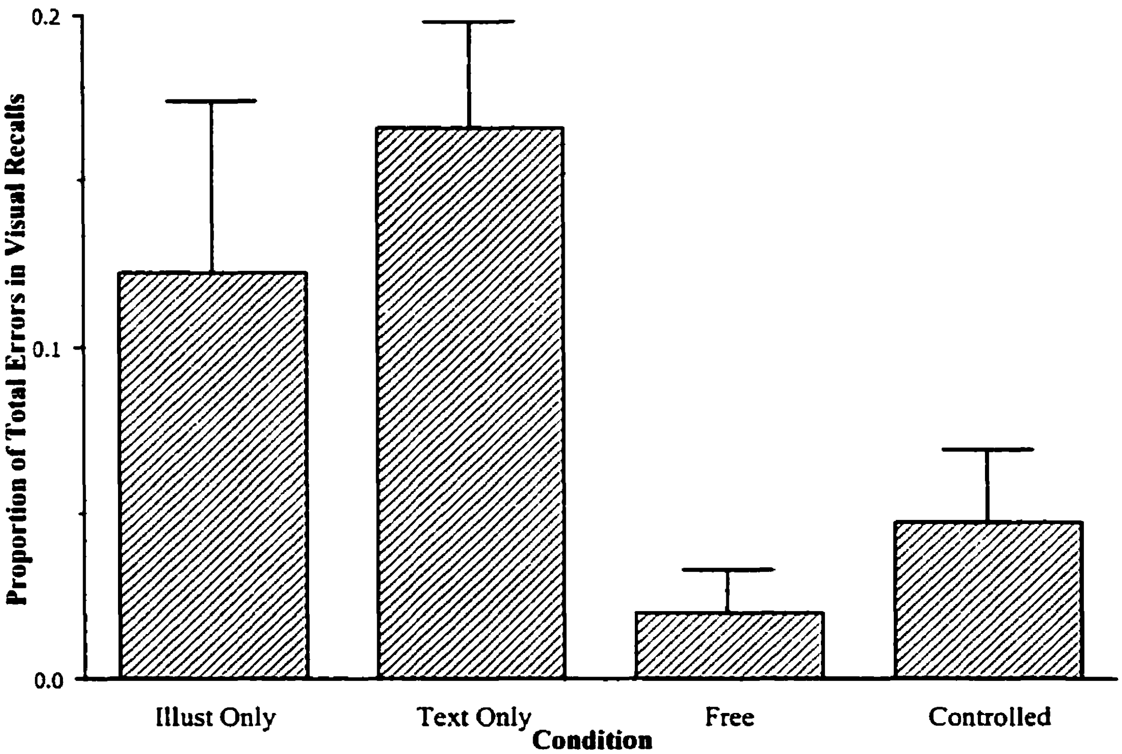


Figure 15. Proportion of errors in visual recall protocols by condition.

Comprehension questions. The final piece of post-input data that was analyzed consisted of participants' responses to eight open-ended comprehension questions. As described in the methods section, these eight questions were collapsed into three types: (a) those that required integration of information presented in the text only (N=3), (b) those that required integration of information presented in both the text and illustrations (N=3), and (c) those that could not be answered directly from the presented materials and required the participant to generate new knowledge (N=2). Examples of participant responses to each question type are available in Appendix E.

Participants' scores for each of the three types of questions were analyzed using a one-way MANOVA with contrasts. The analysis revealed that participants in two text with illustrations conditions did not differ significantly from one another (multivariate $F(3,18) = 0.272$, $p = .845$). The two pooled text with illustrations groups outperformed the text only condition on the set of question types (multivariate $F(3,18) = 3.432$, $p = .039$). Univariate tests revealed that the text with illustrations groups performed significantly better than the text only group on questions that required *integration* of information expressed in the text and illustrations ($F(1,20) = 4.594$, $p = .045$), and in their ability to respond to questions requiring the generation of *new information* ($F(1,20) = 5.172$, $p = .034$). Not surprisingly, the groups exposed to the text with illustrations outperformed the illustrations only group on the set of questions (multivariate $F(3,18) = 10.024$, $p < .0005$). Univariate results were significant for questions requiring integration of information expressed in both text and illustrations ($F(1,20) = 19.095$, $p < .0005$)., questions that required integration of information expressed through text ($F(1,20) = 7.435$, $p = .013$)., and questions that required generation of new information ($F(1,20) = 15.840$, $p = .001$).

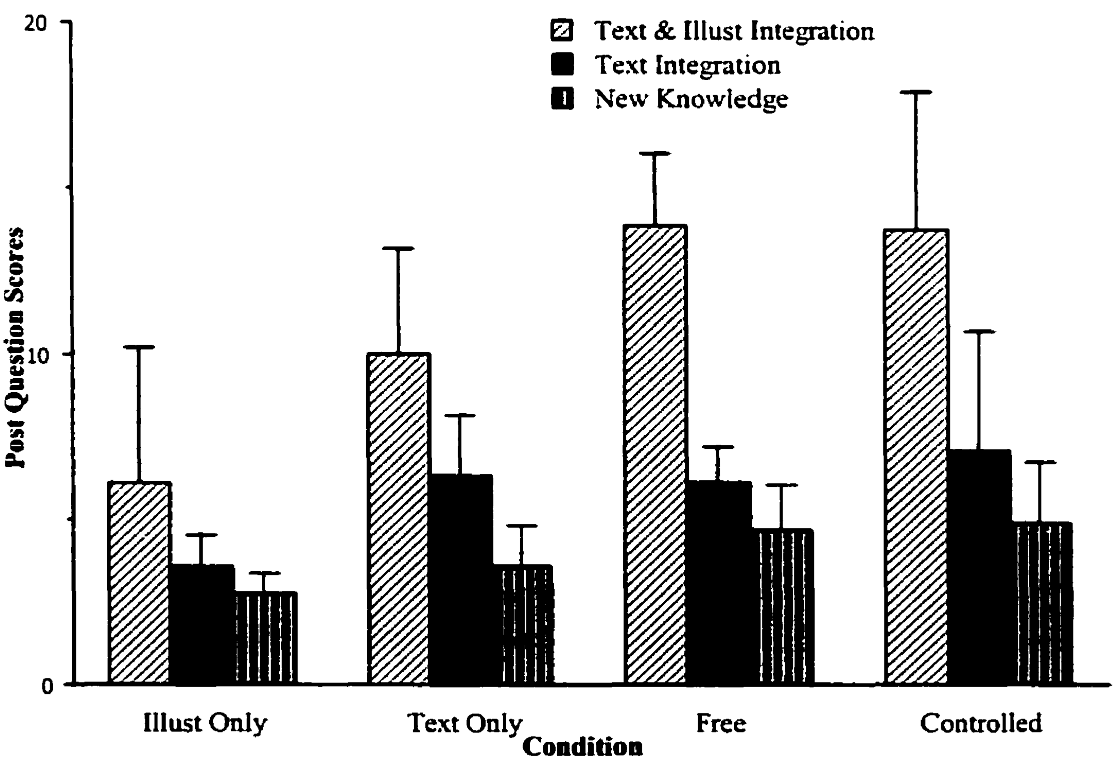


Figure 16. Average scores on post-input comprehension questions by condition by Question type.

Summary of post-input results. The pattern of results obtained for the post-input data indicate that (a) the two groups that were exposed to a combination of text and illustrations did not differ from one another in any measures of their post-input comprehension of the target domain; (b) the pattern of post-input comprehension for the two text with illustrations conditions did differ from the text only condition for non-literal measures of comprehension, and from the illustrations only condition on both literal and non-literal measures; and (c) the groups that were exposed to text differed from the illustrations only condition in terms of the effect of information staging.

Specifically, the two groups that were exposed to text with illustrations (CA and FA) differed from the text only group (TO) in that they demonstrated more elaborated and more connected mental model representations for the type of information that was staged high (i.e., structure information). In addition, the CA and FA groups demonstrated fewer errors in their visual recalls and were more adept than the TO group at answering questions requiring integration of information presented through text and illustrations and questions requiring the generation of new knowledge. Interestingly, the text only group and the text with illustrations groups demonstrated unique patterns of inferences regarding structure information, with the TO group favoring arrangement type inferences and the CA and FA groups favoring location type inferences. Each of the groups that were exposed to text demonstrated a staging effect on several post-input measures reflecting high level comprehension processing. That is, each of these groups demonstrated a preference for structure information in the degree to which they elaborated and integrated concepts at the mental model level and in their visual recall protocols. In contrast, the illustrations only group (IO) focused on both structural and energy information. Finally, the higher level of performance in mental model construction of the two groups exposed to text and illustrations (as compared to the text only group) was largely attributable to differences in their engagement in active on-line processing events (as opposed to interpretative processing events).

Chapter VI: Discussion

The broad research problem addressed in this study was concerned with how individuals are able to comprehend and integrate information within and across modalities in multimedia environments (i.e., how they are able to construct coherent and integrated mental models of the target domain). The main purpose of the current study was to better understand the cognitive processes and representations involved in comprehending and learning from text and illustrations. How do illustrations affect text processing and learning? How does text influence the processing and comprehension of illustration information? How are both media sources used to construct coherent mental models of the domain?

In order to answer these questions multiple measures of comprehension processing were collected as participants read from text, studied illustrations, or learned from a combination of text and illustrations describing the human visual system. Post-input measures were collected to assess overall comprehension and learning. Both the text and illustrations were designed to highlight structure information about the human visual system over functional and energy path information. The data were subsequently analyzed using multivariate and univariate analysis of variance models and log-linear models. The results of these analyses are discussed below. An attempt to integrate the findings of the current study with the existing research literature is made. The contributions of this study to theory and practice are discussed. Finally, limitations of the present study are identified and implications for future research are explored.

Sample Characteristics

Prior domain knowledge. Pre- and post-experimental questionnaires were administered to participants in order to determine each participant's prior knowledge of the human visual system and their experience with text and illustration characteristics similar to those used in the present study. The results of the questionnaires established that the participants were well suited to the purposes of this study. That is, before starting the experimental task, participants revealed scant prior knowledge and formal academic experience with the topic of the human visual system. Following the exposure to the experimental materials, all participants rated themselves as possessing low prior domain-specific knowledge. They were, however experienced in learning from materials possessing similar characteristics to those employed in the study. Based on these findings, it can be concluded that the participants sampled were indeed appropriate candidates for studying learning from text and illustrations in the target domain.

Spatial ability. During the planning phase of this research, it was thought that one characteristic which might affect the outcomes of the study would be the spatial ability of the participants. Therefore, two standardized tests of spatial ability (Paper Folding and Surface Development) were administered to participants at the beginning of the experimental session. Participants' responses on these measures indicated that all were relatively and uniformly high on spatial ability. Statistical tests indicated that the groups assigned randomly to the four presentation conditions did not significantly differ from one another on these ability measures. Furthermore, neither measure of spatial ability correlated significantly with participants' performance on any of the dependent measures of interest. Based on these results it was concluded that the variable of spatial ability could not be characterized as a moderating variable in this study. Spatial ability was therefore ignored in the remaining analyses. It should be pointed out, however, that these results do limit generalizations of the findings of this study to populations of learners who are similarly high in terms of spatial ability.

The research literature relating spatial ability to the ability to learn from graphic displays has reported mixed results. For example, Thorndyke and Stasz (1979) found spatial ability to be positively related to the ability of learners to benefit from training in effective procedures for learning map information. Mayer and Sims (1993) also found spatial ability to be positively related to college student's ability to coordinate visual and verbal representations presented through animation and narration in a multimedia environment. However, there have also been reports to the contrary. For example, Batista (1981) reported no relationship between spatial ability and undergraduate's ability to benefit from a combination of text and graphics as compared to text only in the domain of algebra problem solving with undergraduate students. Further research is required to more carefully investigate the relationship between spatial ability and learning from visual media.

Controlled Vs Free Access to Illustrations

One of the pre-planned contrasts explored in this study involved comparing the comprehension processing of participants who were provided with free access to text and illustrations (FA) with a group of participants who had free access to the text but who received experimentally controlled access to illustration information (CA). Although there were no specific patterns of differences anticipated for this comparison, it was explicitly included to test possible differences between comprehension processing under more natural (FA) verses more experimentally controlled (CA) conditions of access to illustration information. The results of these comparisons revealed no significant differences between these two groups on any of the dependent measures of interest. Based on these findings the two text with illustrations groups were subsequently collapsed and therefore the discussion of the results that follows will not distinguish between these two presentation groups. However, before leaving this issue behind it is first necessary to explore why the two groups did not appear to differ from one another.

One possible reason underlying the failure to find any statistically significant differences between these two groups is insufficient power of the design. It could be argued that six participants per cell is not large enough to obtain sufficient statistical sensitivity to detect true effects if these effects are of small size. Post hoc power analyses were conducted to examine this possibility. The results of the power analyses are displayed below in table 14. As the table indicates, power was indeed low in some cases in which the effects were small. Power values ranged between .052 and .901. Therefore, the possibility that these two groups did differ from one another on some measures remains plausible and needs to be more carefully explored in future research employing larger sample sizes.

Table 14.

Retrospective Power Estimates for Contrast Between FA and CA Conditions.

Dependent Measure		Power		
	Structure	Function	Energy	Multivariate
Reading Time	.487	.371	.053	
On-Line Content	.077	.067	.139	
Verbal Recall	.106	.053	.127	
Mental Model Nodes	.790	.434	.357	
Mental Model Relations	.901	.254	.587	
Visual Recall	.713	.052	.881	
Structure Inferences				.744
Comprehension Questions				.063

An alternative explanation to insufficient power is that participants in both the FA and CA conditions accessed the illustration information in highly similar ways. Exploratory analysis revealed that participants in the free access to illustrations condition (M = 8.83, SD

= 1.47) referred to illustrations only slightly more frequently than participants in the controlled access condition (M = 8)[3]. Thus, in terms of frequency of illustration use these two groups behaved in a highly similar fashion.

The text locations from which participants opted (FA) or were forced (CA) to access illustration information were also highly similar across the two conditions. A descriptive analysis of these data revealed that 30.3% of all the illustration access points occurred at exactly the same text sentence across the FA and CA conditions, and 72.1% occurred within three sentences of each other. Thus, participants within the free condition accessed illustrations in much the same manner as was imposed on participants in the controlled access condition both in terms of frequency and in terms of relative locations within the text. Based upon these behavioral observations one would not expect to see comprehension differences between the controlled and free access conditions in terms of the effect of control over illustration access. In designing the stimulus materials for this experiment great care was taken in deciding where to force points of illustration presentation in the CA condition so that it seemed as natural as possible. Apparently, the experimenter was successful in achieving this goal.

The apparent lack of difference between these two presentation conditions is also related to the issue of learner control. That is, there may be no overall learning benefits associated with providing learners with control of the information flow in learning environments (e.g., hypertext and hypermedia learning environments). The literature comparing the effectiveness of learning from linear text with learning from hypertext displays (in which readers self-determine their path through the text) has reported mixed results (Foltz, 1996). Foltz (1996) has pointed out that the appropriateness of learner control in learning from text may differ depending on such factors as reader goals, prior knowledge, and text type.

Unfortunately, there is insufficient information to distinguish between these possible explanations. It is suggested, therefore, that continuing to distinguish between these two conditions (i.e., free verses controlled access) may remain important in future research, especially since it is possible that either one of these forms of access to multimedia learning environments may be a more effective way of promoting comprehension and learning, and may differ depending on characteristics of the learner, task, materials, and context.

[3] The number and location of illustration access points was held constant across individuals within the CA condition.

Review of Hypotheses

Information staging. At the outset of this experiment several different hypotheses were formulated. Specifically, it was expected that comprehension processing devoted to structure information would be privileged over the processing of function or energy information for all groups due to the staging manipulation. The rationale underlying this hypothesis is a logical extension of past research documenting the effect of information staging on text processing (e.g., Britton, Glynn, Meyer, & Penland, 1982; Clements, 1979; Kieras, 1981; Marshall & Glock, 1978) to it's effects on the comprehension processing of illustrations and text. From a theoretical point of view, this hypothesis implies that information staging is an important dimension of both text and illustrations and that this dimension significantly influences comprehension processing.

In the context of the current study, evidence of information staging effects should be manifest in a pattern of longer reading times, more frequent text look-backs, proportionally more on-line interpretations, greater literal recall in both verbal and visual modalities, more inferences, and more elaborated and connected mental models for *structure* information over and above *energy* or *function* information.

Presentation conditions. The second hypothesis tested was that the groups exposed to a combination of text with illustrations (FA and CA) would demonstrate qualitatively different on-line processing and superior high level understanding of the information that was staged high than the text only group (TO). Finally, the third hypothesis tested concerned the expectation that the groups exposed to a combination of text with illustrations (FA and CA) would demonstrate qualitatively different on-line processing and superior overall comprehension of all information than the illustrations only group (IO).

The rationale underlying both of these hypotheses follows from a model of comprehension processes for both text and illustrations that involves the active construction of representations at multiple levels. In the multilevel comprehension model, the learner attempts to build integrated representations of the target domain. Furthermore, within the context of this model it was hypothesized that text and illustrations each would contribute to comprehension processing in different ways. Specifically, it was hypothesized that text would provide a means for elaborating the semantic interpretation of illustration information and that illustrations would facilitate conceptual processing at the mental model level by providing an explicit external context for framing the comprehension of relevant text information. From a theoretical point of view, this hypothesis assumes that cognitive models of discourse comprehension can be extended to the case of comprehending and learning from illustrations and text.

In the context of the current study, evidence that illustrations facilitate text processing at the mental model level should be manifest in a pattern of more elaborate and integrated mental models for groups that were exposed to both media (FA and CA) than for participants in the text only group (TO). The expectation that these effects should be limited to the type of information that is highlighted follows from hypothesis one above. Evidence that text facilitates overall comprehension processing should be manifest in a general pattern of superior learning performance for the text with illustration groups (FA and CA) than the illustrations only group (IO).

Discussion and Interpretation of Results

Several different on-line measures were collected in order to obtain data that were temporally connected to participants' processing and on-going comprehension of the materials. The on-line measures collected consisted of (a) sentence reading times for groups that were exposed to text (FA, CA, and TO), (b) traces of text look-backs for all groups that were exposed to the text (FA, CA, and TO), and (c) on-line interpretation protocols for all groups. In addition, several different post-input measures were collected to assess the effects of different presentation conditions and information staging on comprehension and learning. Three different post-input measures were collected for all groups including (a) verbal recall protocols, (b) visual recall protocols, and (c) answers to free-response comprehension questions. The results obtained for each of these measures are discussed below in terms of the experimental hypotheses.

The effect of information staging. The finding that reading times were significantly longer on average for structural than for functional and energy information reflects a classic staging effect (e.g., Clements, 1979; Kieras, 1981) and supports the hypothesis that readers devote more comprehension time and resources to processing information from the perspective that is privileged or signaled as important in the text structure. Analysis of text look-backs that participants made during the reading task also indicated a main effect of information staging where all text groups made significantly more text look-backs involving structure information than they did for function or energy information. In addition, the analysis of participants' on-line protocols revealed similar staging effects in that participants who were exposed to the text (i.e., TO, FA and CA) devoted proportionally more of their on-line statements to describing structure information than either function or energy information. Again this provides evidence of the special importance readers devoted to processing information from the perspective that was staged high. Participants in the illustrations only (IO) condition, however, did not show the anticipated staging effect. Rather, the results indicate that these individuals were equally focused on both structure and energy path information.

When the measure of propositional recall was considered, there was no evidence of information staging effects for any of the conditions. That is, the probability of verbally recalling different types of information did not depend on how that information was staged in the text perspective. This finding suggests that perspective staging effects may not be evident at the local propositional level of text processing. The failure to find such an effect on this measure may not be surprising since information staging was controlled at a high conceptual level (i.e., information perspective) rather than a low topic level. Taking this position, it may make more sense to expect to find information staging effects to be evident at higher levels of comprehension and learning.

The visual recall data, however, did provide support for the information staging hypothesis. All participants who were exposed to the text (FA, CA, and TO) provided proportionally more structure information than function or energy information in their visual protocols. Furthermore, this staging effect was larger for groups that were exposed to both text and illustrations (FA and CA) than for the TO group, providing further support for the hypothesis that exposure to two media enhanced the staging effect. It may be that visual recall requires higher levels of conceptual processing that is more sensitive to the staging manipulation than does recall of verbally presented propositions. Again, the illustration only (IO) group demonstrated a unique pattern of recall with equal preference given to both structure and energy information, rather than to structure information alone. This specific pattern was not anticipated and suggests that the staging manipulation was not effective for the IO group.

In general, the pattern of on-line findings supports the hypothesis that readers devote more processing effort to understanding information that was staged high (i.e., structure information) for all the groups that were exposed to text (TO, FA, and CA). Furthermore, this effect was increased through the addition of illustration information. This pattern did not carry over to participants in the illustrations only group. Rather, the on-line interpretations for the IO group indicate a focus on both structure and energy information. This result was unexpected and interpreted as suggesting that information staging effects may be particular to text.

The post-input data provided partial support for the information staging hypothesis. All the conditions in which participants were exposed to text (CA, FA, and TO) demonstrated preferential recall for information of the type that was staged high (i.e., structure) in their visual, but not in their verbal recall protocols. On the other hand, participants in the IO condition failed to demonstrate the anticipated staging effect at all. Again, this finding seems to indicate that the staging effect was carried through the text and not through the illustrations.

Such findings have clear implications for the design of illustrated instructional texts. Specifically, these findings indicate one way in which the structure of text can be intentionally manipulated to signal to the learner what conceptual perspective to regard as "important".

The effect of illustrations on text processing. Analysis of sentence processing times indicated that participants with access to both text and illustrations (CA and FA) read the text sentences significantly *faster* than participants in the text only condition (TO), regardless of information type. Thus, the presence of illustrations in the context of text led to an overall decrease in text processing time. Additional analysis, however, revealed that the conditions did not differ in terms of overall learning time. That is, when the time readers devoted to processing illustrations was considered in addition to text processing time, the CA and FA conditions did not differ from the TO condition.

This set of findings does not support a bottom-up account of multimedia comprehension processing in which it would be expected that readers who were exposed to both media would demonstrate longer reading times than the text only group. The rationale underlying such a position is that such readers would face the task of comprehending text information, illustration information, and of integrating information across the two modalities. Rather, these findings support a more top-down account of multimedia learning where the presence of illustrations acts like a context or prior knowledge effect on text processing. According to this type of account, the presence of illustrations may *constrain* text processing by providing readers with an external representation or frame that contains aspects of the conceptual knowledge representation that needs to be generated for effective comprehension in the absence of illustrations. That is, readers in the text only condition may have had to perform additional inferences on the basis of the text information in order to achieve comprehension results that are similar to those of participants who had access to both text and illustration information. These inferences may be more difficult to make when the visual representation is lacking. This idea fits well with similar arguments put forward by Winn (1987). Specifically, he has claimed that illustrations may decrease working memory demands and thereby enable the learner's limited cognitive resources to be devoted to higher order operations such as developing a coherent semantic macrostructure. Other researchers have constructed similar arguments. For example, Bauer and Johnson-Laird (1993), Larkin and Simon (1987), and Zhang (1997) have claimed that information presented through external visual representations may positively bias problem-solving performance since they present related information simultaneously and can thereby be used to directly support the construction of inferences that may be relevant to successful performance.

Further support for the position that participants in the text only condition may have experienced a more cognitively demanding comprehension task than individuals in the combined text and illustrations conditions is indicated by the results of the text look-back analysis. Analysis of these data indicated that the text only group performed longer look-backs that involved traversing greater text distances than did the text with illustrations groups. This was particularly true for structure information. Taken together, these results imply that participants in the TO group were attempting to integrate disparately presented text information of the type that was staged high as they were reading, but had to reread large sections of the text to do so.

In terms of the on-line protocol data, the groups who were exposed to both text and illustrations provided significantly more on-line interpretations describing structural information than those who were exposed to the text only. This finding indicates that the presence of illustration information in the context of text was associated with larger staging effects than were observed in the text only condition.

Participants' on-line interpretations were also coded for the type of comprehension processing event that was concurrently demonstrated. Log-linear analysis of these data indicated qualitatively different patterns of on-line processing across conditions. Essentially, participants who were exposed to text with illustrations demonstrated a much more active profile of on-line interpretations than did those in the text only group. In particular, readers who were exposed to both sources of information demonstrated more use of prior knowledge, metacognition, strategies, and information integration than did participants who were exposed to text alone. This effect was particularly large for on-line interpretations in which structural information was described. Thus, both the text only (TO) and combined text and illustrations conditions demonstrated evidence of on-line information integration. The text only group demonstrated evidence of attempts to integrate the text information vis-à-vis their look back behavior. On the other hand, the two text with illustrations conditions demonstrated evidence of information integration by way of their on-line protocol data.

The text only (TO) and the text with illustrations (FA and CA) groups did not differ from one another in terms of their literal comprehension of the materials as expressed through either verbal or visual modalities. However, the TO condition did commit more errors in their visual recall protocols. Generally speaking, however, the presence of illustrations in the context of text did not significantly influence readers' ability to comprehend information at the propositional level of representation.

An analysis of the types of structural inferences did indicate significant differences between the conditions. Individuals in the TO condition made significantly more inferences regarding the *arrangement* of components in the visual system, whereas, individuals in the two text with illustrations conditions made significantly more inferences regarding the *relative location* of components. This pattern of findings suggests that text and illustrations support qualitatively different types of inferences.

Participants' verbal descriptions of the visual system were coded against three different mental models that correspond to high level conceptual representations of the components of the human visual system organized according to (a) structural, (b) functional, and (c) energy path information. Analysis of these data revealed that participants who were exposed to both text and illustration information (FA and CA) generated more *complete* and more *integrated* representations for the type of information that was staged high (i.e., *structure*) as compared to participants in the text only (TO) condition. These results provide support for the hypothesis that illustrations specifically facilitate comprehension at the mental model level of representation and that the facilitation effects are sensitive to the information perspective presented. Glenberg and Langston (1992) have reported similar results where they found that the addition of illustrations to text affected comprehension at the mental model level but not at the literal level of comprehension.

The results of the post-question data also support the hypothesis that a combination of text and illustrations facilitates the development of an integrated and flexible knowledge structure. Specifically, the results indicated that individuals in the combined text and illustrations conditions were more adept at answering questions that required theintegration of information presented through text and illustrations and questions that required the learner to generate new information. Mayer and his colleagues have also reported a similar finding. Specifically, Mayer has reported evidence that adding explanative illustrations to text helps low prior knowledge readers develop flexible knowledge (e.g., Mayer, 1993, 1997; Mayer & Gallini, 1990).

In general, the pattern of on-line findings indicates that access to illustration information significantly affects text processing. That is, access to both text and illustration information not only sped up text processing in general, but also actually supported a more active profile of on-line processing of information that was signaled high in the text structure. Thus, the hypothesis that participants exposed to a combination of text with illustrations (FA and CA) would demonstrate qualitatively different on-line processing of the information that was staged high than the text only group (TO) was supported by the on-line results.

While the text only group demonstrated slower text processing and less active on-line interpretations, the results of the text look-back analysis indicates that these individuals were also engaging in attempts to integrate text sentences describing structure information. The fact that this group demonstrated far fewer instances of integrative processing in their on-line protocols and that the structural mental models they constructed were less developed and integrated than those of participants who were exposed to both media indicates that they were not as successful in this regard as the groups exposed to a combination of text and illustrations.

Taken together, these results support the position that adding relevant illustrations to text descriptions specifically benefits the ability of low prior knowledge learners to generate high level, integrated, and flexible knowledge structures. Furthermore, this facilitation effect was sensitive to the conceptual perspective from which the information was described.

The effect of text on illustration processing. Analysis of participants' on-line protocols indicated that participants in the illustrations only (IO) condition demonstrated a different pattern of on-line interpretation from the remaining conditions in which they placed emphasis on processing *both* structure and energy information. Thus, the nature of the information staging effect was different for participants who were only exposed to illustration information than it was for conditions that were exposed to both media (CA and FA). This finding suggests that the information staging effect was not demonstrated in the illustration only condition and may have depended on the presence of text.

Analysis of the on-line processing events also demonstrated that the pooled CA and FA conditions made proportionally more inferences, with more integration, more strategy use, and more metacognition in their on-line processing of the materials than the IO group. Again, exposure to both media seems to have enabled learners to be far more active in their processing of the content than exposure to a single information medium.

The text with illustrations (FA and CA) groups outperformed the illustrations only (IO) group in terms of their ability to recall structure and function propositional information in both verbal and visual modalities, but did not differ in their ability to recall energy propositional information. An analysis of the errors made in the visual protocols revealed that individual's exposed to a combination of text and illustrations made fewer errors than either the text only or illustrations only group. Thus, the hypothesis that the combined media conditions would demonstrate better learning outcomes that the IO condition was partially supported by the verbal and visual recall data. Interestingly, further descriptive analysis of the IO protocols suggested that these individuals were using energy transformation information to organize their verbal recall of information.

In addition, students in the FA and CA conditions developed more elaborate and integrated mental models than the illustrations only (IO) group for structure and function information. The groups, however, did not differ in terms of their ability to elaborate mental models of energy information. Again, these data suggest that while the FA and CA groups were generally more successful than the IO group in terms of comprehending and learning the domain, the IO group was just as successful when dealing with energy information.

Finally, the two pooled FA and CA conditions outperformed the IO condition in terms of their ability to answer each type of post-comprehension question. Thus, learning from a combination of text and illustrations led to the development of more integrated and flexible knowledge structures than did learning from illustrations alone.

In general, the on-line results support the position that text affects illustration processing. That is, the text with illustrations group demonstrated a different pattern of comprehending content and used a different pattern of comprehension processing than did the illustrations only group. Thus, the expectation that the groups exposed to a combination of text with illustrations (FA and CA) would demonstrate qualitatively different on-line processing than the illustrations only condition is supported by these results. Furthermore, the pattern of post-input results provided partial support for the hypothesis that exposure to a combination of text and illustrations generally facilitates comprehension and learning as compared to learning from illustrations only. The exception to this conclusion concerned the ability to comprehend and learn energy path information. Generally speaking, participants in the illustrations only group did not differ from the participants in the combined media conditions in their overall comprehension of energy path information. These results suggest that the energy path information was used as a perspective from which to understand the illustration information for these individuals. This preference may relate to how engineering students view energy transformation information in systems of this kind.

A link between on-line processing and comprehension outcomes. Two consistent variables which distinguished the combined media groups (FA and CA) from the conditions that were exposed to a single information medium (TO and IO) were (a) active on-line processing episodes, and (b) mental model development. A MANCOVA was performed to investigate whether the pattern of more *active* on-line processing was responsible for the observed condition effects on the ability to describe more elaborated and integrated mental models. This analysis revealed that when active on-line processing was used as a covariate, the significant differences between presentation conditions that was observed for the mental model elaboration variable (i.e., *nodes*) were attenuated. Furthermore, similar analysis on the mental model integration measure (i.e., *relations*) indicated that the condition differences were partially attenuated when the active on-line processing variable effects were partialled out.

These results indicate that the degree to which participants were "active" in their on-line processing of the learning materials was at least partially responsible for differences in learning outcomes. Thus, the text with illustrations conditions led the students to employ more active processes in mental model construction, and these processes enabled them to construct more elaborated and connected (i.e., coherent) mental models.

General Discussion

In general, the pattern of post-input results indicates that overall comprehension is benefited in several ways by a combination of text and illustrations. In particular, exposure to both media facilitated the construction of high level mental model representations, and ultimately the development of more flexible knowledge structures when compared with the text only group. When compared with the illustrations only condition, the combination of text and illustrations facilitated performance on all post-input measures (with the exception of visual recall).

The pattern of facilitation effects described above, however, was different for various information types. That is, the facilitation of mental model construction that was observed for the contrast to the text only group was particular to *structure* information. On the other hand, the more general facilitation effects observed for the contrast to the illustration only group was limited to structure and function information only.

Such results suggest that the ability to comprehend each information type (structure, function, and energy path) may depend on characteristics of the media. That is, one potential positive bias of text, when compared with illustrations is that text favors processing of relational semantic information. That is, particular relationships between concepts can be directly encoded in the text, whereas, relational information encoded in illustrations is left implicit. Furthermore, the results of the current study also suggest that staging effects may be a second potential bias of text. On the other hand, it could be argued that illustrations may be advantageous over text descriptions in that they are able to present spatially and structurally related information simultaneously to the learner. How do such differences affect comprehension processing of each information perspective?

The current results suggest that the addition of illustrations (i.e., the property of simultaneous information presentation) to text specifically facilitated high level comprehension of structure information. Furthermore, the addition of text (i.e., the properties of staging perspective and explicit relations among concepts) facilitated both local and high level comprehension of structure and function information. Thus, it appears that both text and illustrations are particularly beneficial to the comprehension of structural information about a functional system. In contrast, illustrations do not appear to be an effective means for learning about the function information, as the illustrations only group demonstrated poor understanding of this aspect. The comprehension of energy information in general appears to be equally understood from either media as there were few differences between the conditions for this type of information.

As with the on-line results, the post-input results also provide evidence that: (a) text influenced illustration processing, and (b) illustrations in turn affected text processing. Support for the first claim is evident from the observation that the staging manipulation was larger in the text with illustrations groups than in the text only group, while there was no observed staging effect for the illustrations only group. The focus on structure in the text appears to have shifted the conceptual processing of the illustrations away from an emphasis on energy path information (as was evident in the illustrations only condition). Support for the second claim is indicated by the finding that adding illustrations to the text improved mental model construction and the ability to answer free-response comprehension questions as compared to the text only condition. This supports the idea that individuals actively attempt to integrate information from text and illustrations rather than from two separate representations.

The finding that a combination of text and illustrations supported processing of privileged (i.e., structure) information specifically at the mental model level of comprehension needs to be explained. How was it that these individuals were able to construct a more elaborated and connected understanding of the targeted information? The fact that the proportion of active on-line processing was positively related to the development of more elaborated and integrated mental models suggests a direct link between these differences in on-line processing and comprehension at the mental model level of representation.

Both the on-line and post-input results suggest that the signal for information staging was specifically carried through the text structure. Thus all participants who were exposed to text devoted more processing effort to the comprehension of structural information. With the addition of illustration information, participants were also able to devote more active processing to this targeted information type. That is, during the information presentation phase of the experiment, these individuals demonstrated more information integration, greater use of comprehension strategies, etc. This more active processing was subsequently associated with the development of more elaborate and integrated mental model representations of structure information.

The research literature has presented similar findings. For example, Chi, de Leeuw, Chiu, and LaVancher (1994) have reported evidence of a link between the propensity to generate self-explanations while reading, the ability to integrate text-derived information with prior knowledge, and the ability to construct appropriate mental model representations. Peeck (1993) has argued that large learning gains could be made by encouraging readers to *actively* manipulate illustration information. Similarly, McNamara, Kintsch, Butler Songer, and Kintsch (1996) have also provided evidence that active processing is associated with comprehension benefits at the situation model level rather than at the text-base level of understanding. More recently, Davidson-Shivers, Shorter, and Jordan (1999) employed the think aloud technique to study the learning strategies and navigational decisions that fifth grade students made while learning a Hypermedia lesson. These researchers found that the group of students who scored highest on a post-test had demonstrated a greater variety of learning strategies on-line than did students who scored either in the average or low post-test score ranges. In addition, while the low ability students did show evidence of on-line constructive processing, their protocols tended to contain conceptual errors. Taken together, such findings coincide well with the results of the current study in supporting the idea that active on-line processing is associated with positive learning outcomes, particularly at high levels of comprehension. The results of the current study further demonstrate that one way to engage readers in a pattern of more active processing is to provide them with alternative but related external representations (e.g., illustrations).

Conclusions

The effect of information staging. The effect of information staging appears to be carried through the text (which provides a perspective on the information) and *not* through the illustrations. However, when text and illustration information are both available, the effect of information staging is stronger than in the case of text only. These findings suggest that in the context of illustrations, the text served the important function of providing readers with signals as to what information perspective is "important" and thus helped provide a coherent and consistent structure from which to understand the target domain.

The fact that participants in the illustrations only condition did not demonstrate the anticipated staging effect requires some elaboration. It appears that these individuals tended to focus on both structure and energy path information (i.e., both information types were directly available through the illustrations), and used the energy path information to provide a conceptual framework for their interpretation and recall of structure information. Further descriptive analysis suggested that the energy path information provided these individuals with a narrative schema into which they could incorporate their descriptions of structure information. Perhaps in the absence of an elaborated linear structure (i.e., text) it was difficult for these individuals to interpret structural information in a systematic way. That is, structural information as presented though illustrations can be interpreted and related from many perspectives (i.e., left to right, right to left, outside to inside, on top to underneath, etc.). On the other hand, the depiction of energy path information may have provided a more constrained and coherent framework from which to understand the domain. This visual description provided path information (indicated through lines and arrows) that includes a sequence of transformations of light (i.e., a sequence of events) that are produced by functional components of the visual system which are arranged in a certain structure. Thus, the depiction of energy path information contained information about both spatial relations and temporal relations, whereas the depiction of structural information only contained spatial relation information. Therefore, in this case, individuals in the illustrations only condition may have opted to interpret the illustration information from the energy path perspective which provided them with a more coherent (and constrained) framework from which to understand the human visual system. This may have been particularly effective for engineers.

This idea fits well with past research in text processing which demonstrates that placing text within a well known narrative event schema can improve comprehension (e.g., Poulsen, Kintsch, Kintsch, & Premack, 1979). From a semantic perspective, narrative frames are well defined and more constrained than other types of high level conceptual frames, including descriptive frames (e.g., Denhière, & Denis, 1989). Adopting this position, a plausible explanation of these results is that the text provided sufficient elaboration of the structure of the human visual system with adequate coherence signals to enable readers to construct an organized representation of that information. It would be interesting to examine in future research whether the effects of organizing the text as a narrative description of transformation of the light energy would have led to different comprehension outcomes.

This explanation suggests that discourse can play an important role in directing learners' comprehension processing of illustrated information in particular, and perhaps other external sources of information in the learning situation as well. Theorists in the situated learning camp have made similar claims (e.g., Greeno, 1989; 1991) by taking the position that discourse provides a major means by which diffuse information sources in the learning situation can be elaborated, related, and integrated. Thus, discourse may play a privileged position in providing a "framing" perspective from which to understand related sources of information.

How illustrations affect text processing. Based on the pattern of results observed in the present study, it is possible to make several conclusions with regard to the question of how illustrations can affect text processing. First, illustrations specifically support comprehension at the mental model level of representation. This finding confirms the results of previous research presented in the literature (e.g., Mayer, 1993, 1997; Hegarty & Just, 1989, 1993). Such findings have direct implications for the design of learning materials. Clearly, for low prior knowledge readers faced with the task of understanding and learning about a functional system, the addition of illustrations relevant to the text is desirable.

The question of how illustrations facilitate the construction and elaboration of mental model representations can also be addressed. The current study provides evidence that illustrations provide a form of external representation that helps readers construct a coherent and integrated conceptual understanding of the text information. Thus, presenting text in the context of illustrations has the potential to reduce some of the reader's processing load in interpreting the propositions by providing top-down constraints on comprehension processing. Thus, readers are able to become more active top-down processors in their comprehension of the materials. This pattern of findings suggest that illustrations are beneficial because they provide a specific external context within which text information can be understood and integrated across sentence and paragraph boundaries, and with the participant's prior knowledge to form a high level conceptual understanding.

In addition, there was evidence that exposure to illustrations was associated with the use of a different pattern of inferences than was found in the text only group. It appears as though text and illustrations each supported qualitatively different types of inferences. The question of precisely how such inferences relate to differences in overall comprehension as well as how they were related to specific aspects of the target content will require further research.

How text affects illustration processing. The results of this study also support specific conclusions regarding the question of how text affects illustration processing. Given the manner in which the materials were designed for the present study, it is not surprising that learning from illustrations alone was not as effective as learning from text only or a combination of text and illustrations. The text provided much more information (and more information types) than the set of illustrations. Thus, in the current study text served the function of elaborating illustration information. However, beyond this finding, there is also evidence that the staging effect was specifically controlled through the text. That is, in the absence of text, participants in the illustrations only condition relied principally on energy information to provide a high level structure for interpreting and integrating structural descriptions of the visual system even though an attempt was made to emphasize or privilege the structural perspective in the illustrations. It may well be that when learning from illustrations alone, energy path information provides a better perspective from which to understand a functional system like the human eye.

Again, there was evidence that exposure to the text was associated with the use of a different pattern of inferences than was found in the groups with illustrations. Future research will be required to more closely investigate the theoretical and practical significance of this finding.

Multimedia processing. The current results suggest that each medium influences the processing of the other. That is text affects illustration processing and vise versa. Specifically, each form can constrain comprehension processing in unique ways. Text can provide constraints on illustration processing by providing a signaling system for focusing on some information over other information (i.e., a conceptually based staging perspective). On the other hand, illustrations can constrain text processing by facilitating information integration in a specific externalized context. Such findings suggest that it is precisely this quality of mutually constraining information sources that may underlie the potential effectiveness of multimedia learning environments. These mutual constraints favor the construction and elaboration of integrated and coherent mental models.

Such results suggest that the study of learning in multimedia environments may be more appropriately approached from a perspective which explicitly addresses the potential reciprocal effects that each medium may place on the others. That is, learners in such situations interpret multiple sources of information within the learning context as related to one another rather than as separate entities. Clearly, learning from a combination of text and illustrations is not the same as learning from text plus learning from illustrations. Rather, each source of information contributes to overall comprehension and learning by providing potential constraints on constructing a conceptual understanding the domain.

Learning in Context. The finding that comprehension processing of the text and illustrations each affected one another suggests that a text and aspects of its nonlinguistic context (i.e., illustrations) are not understood separately from one another. Rather, the learner's construction of *a model of the situation* (i.e., a mental model) is shaped by *both* linguistic and nonlinguistic information. Such findings argue for a need to extend cognitive models of discourse processing to consideration of features beyond the text and the reader to an explicit consideration of the learning context - including illustrations.

By the same token, the study of "leaning in context" or "learning from a situated point of view" requires sensitivity to and the ability to define what the relevant sources of information are, what content or meaning they encode, how the learner is able to construct meaning for each source, and how the learner is able to integrate information across sources.

Methodological concerns. The current study provides a clear demonstration of the importance of collecting multiple indices of comprehension processing that span different times in processing. This clearly follows from the view that comprehension is a multilevel and multi-layered process. Each of the measures collected in this experiment was found to provide unique information about participants' comprehension processing. By integrating the analysis of all these measures, a coherent account emerged of how illustrations and text each contributed to overall comprehension and learning. A very different account would have resulted if fewer measures had been employed. For example, without the information provided by the on-line protocol data, it would not have been clear how the presence of illustrations affected the on-line construction of mental model representations. The implications for studying a complex problem such as comprehension and learning in multimedia environments are evident.

Contributions and Implications

The research reported here makes several contributions to the existing literature on comprehending and learning from illustrated text. First , the attempt to obtain data reflecting the content of readers' on-line processing and their evolving comprehension of text, illustrations, and illustrated text at multiple points in processing and the analysis of these data in terms of the levels and kinds of semantic representations that are affected is unique. In particular, such data yielded important information about the content of low prior knowledge readers' evolving understanding of each medium and of both media. This information has led to clarification with regard to the mechanisms underlying the facilitative effects of illustrated text on comprehension processing.

Second, the finding that a combination of text and relevant illustrations was associated with more robust comprehension and learning of the target domain clearly indicates the value of adding relevant illustrations to text. Specifically, the results imply that illustrated text is beneficial for low prior knowledge learners faced with the task of learning about a complex functional system. The results also support the general notion that learning in multimedia environments may be more effective at fostering a high level conceptual understanding of the target domain than learning through a single medium.

Third, this study has provided evidence of a direct link between learning outcomes (mental model comprehension) and a specific process measure (i.e., active on-line processing of content). This finding clearly indicates that at least one of the ways in which illustrations facilitate learning in the context of text information is that they enable learners to engage in more active processing as they are learning.

Fourth, this study has provided evidence that illustrations affect text processing, and vice versa. Specifically, the way in which illustrations affect text processing is by enabling the reader to be more active in the processing of content, which in turn is associated with the ability to form more elaborated, connected, high level, and flexible mental model representations. On the other hand, text affects illustration processing by providing a means for elaborating the semantics of illustration information, and a means for signaling important information through staging. This suggests that consideration of text structure needs to be considered when adding illustrations to learning materials. That is, features of the text (i.e., staging) in which illustration information is embedded can have significant influence on how illustration information is processed. More generally, such findings support the idea that there is a need to extend cognitive models of discourse processing to consideration of features beyond the text and the reader to an explicit consideration of the learning context.

Finally, the study reported here indicates that models and methods from discourse processing research can be successfully applied to the study of learning in multimedia environments. In the introduction to this research problem the popular movement towards multimedia learning environments was noted. It was also noted that our current understanding of how learning takes place in such environments is poor. Part of this problem stems from an apparent lack of methods and models for studying learning in such complex environments. The current study has provided a clear demonstration of the utility and ability to extend existing cognitive models of discourse processing to multimedia research problems.

Limitations and Future Research

The study of comprehension processing and learning in multimedia environments is a complex research problem. Clearly, all the relevant issues involved in this area cannot be addressed in any single study. Rather, an entire research program is required to adequately address the range of questions and issues involved in understanding learning in multimedia environments. In an attempt to simplify this complexity, the current study was limited in several ways that need to be explicitly identified.

First, the population characteristics of the particular students who participated in this study need to be considered when attempting to generalize the findings. That is, the participants of this study were selected to possess low domain specific knowledge, yet were familiar with the types of text and illustrations used in the study. How these results would generalize to individuals with differing amounts of domain knowledge or varying amounts of experience with the material types is an open question that requires further research. An interesting study for future research would be to sample low prior knowledge learners from different domains. For example, biology students may be more biased towards a structural or functional perspective in terms of how they comprehend and learn about a functional system, whereas electrical engineers may be more biased towards understanding such a system from an energy path point of view.

In addition, it is not known how the results of this study would generalize to comprehension processing and learning in other types of domains, utilizing different types of text or, illustrations. For example, whether or not similar findings would recur in a study on learning and comprehending a text with illustrations describing how electrical circuits function (which also have structural, functional, and energy path information) is not known.

A further constraint on generalizing these results is that the participants of this study were of high spatial ability. Future research should be concerned with how such ability variables may influence the way in which learners process and learn from different media.

In addition to population and task considerations, a further limitation of the current study is that the staging of information was controlled but not manipulated independently of information type. Future research in which information staging is directly manipulated for different information types would certainly be a worthwhile endeavor and would assist in establishing the generality of the results reported here. For example, a replication of this study in which energy path information was staged high would be interesting given the preference for such an organization among learners in the illustrations only condition. On a related note, an interesting study to conduct would be to investigate how tutors (maybe from different domain perspectives) naturally stage their descriptions of the human visual system through their discourse when tutoring students, and the effects that this would have on students' comprehension and learning of the domain. The results of such a study would make an interesting comparison to the results obtained here.

Yet another limitation of the current study is that only two different media were considered (i.e., text and illustrations). Whether or not similar findings would result from learning environments comprised of more than two representational media, or whether the effects would be larger is also an open question and requires further research. If learning naturally occurs in multimedia environments, how is it that learners are able to effectively navigate and use information from multiple sources? This whole area of investigation is neglected if research focuses on comprehension and learning from a single media source such as text or illustrations.

Finally, the failure to find any robust staging effects for the illustration only group may have been limited by the way in which information staging was manipulated in the current study. That is, color and brightness were used as cues to visually represent the dimension of information staging in the current study. Past researchers have indicated that color might be a good cue for signaling information cohesion in visuals (e.g., Dwyer, 1972; 1978), however, it is possible that alternative ways of signaling staging may be required to communicate importance and relatedness to learners using visual representations (e.g., physical proximity, similarity, animation, etc.). It is also possible that the free response measures employed in the current study (i.e., free recall and free response questions) may have under-estimated the IO group's comprehension of function information. That is, these individuals may have opted to frame their understanding of functional aspects in terms of energy path descriptions (i.e., one of the types of information that was explicitly available to them) . A more directed assessment of their understanding (e.g., using a probing task) of the functional aspects of the human eye may have resulted in a more sensitive assessment of their understanding of this perspective. Clearly, additional research is needed on the understanding and interpretation of visual representations of information. In particular, we need more information about the inherent biases (both advantageous and disadvantageous) of different representational media for communicating conceptual information to learners.

References

Alesandrini, K L. (1984). Pictures and adult learning. Instructional Science, 13, 63-77.

Anderson, R. C. (1993). The future of reading research. In A. P. Sweet, & J. I. Anderson (Eds.), Reading research into the year 2000 (pp. 17-36). Hillsdale, NJ: LEA.

Anderson, R. C., & Armbruster, B. B. (1984). Content area textbooks. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), Learning to read in American schools (pp. 193-224). Hillsdale, NJ: Erbaum.

Anglin, G. J. (1986). Prose-relevant pictures and older learners' recall of written prose. Educational Communication and Technology, 3, 131-136.

Arens, Y., Hovy, E., & Vossers, M. (1993). On the knowledge underlying multimedia presentations. In M. T. Maybury (Ed.), Intelligent multimedia interfaces (pp. 280-306). Menlo Park, CA: MIT Press.

Baggett, P. (1975). Memory for explicit and implicit information in picture stories. Jounral of Vebal Learning and Verbal Behavior, 14, 538-548.

Baggett, P. (1989). Understanding visual and verbal messages. In H. Mandl, & J. R. Levin, (Eds.), Knowledge acquisition from text and pictures (pp. 101-124). Amsterdam: Elsevier.

Batista, M. (1981). The interaction between tow instructional treatments of algebraic structures and spatial-visualization ability. Journal of Educational Research, 74, 337-341.

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. Psychological Science, 4, 372-378.

Beveridge, M., & Griffiths,V. (1987). The effect of pictures on the reading proceses of less able readers: A miscue analysis approach. Journal of Research in Reading, 10, 29-42.

Bieger, G. R., & Glock, M. D. (1985). The information content of picture-text instructions. Journal of Experimental Education, 53, 68-76.

Bieger, G. R., & Glock, M. D. (1986). Comprehending spatial and contextual information in picture-text instructions. Journal of Experimental Education, 54, 181-188.

Bluth, L. F. (1981). The illustration-comprehension relationship: A review of the literature. Reading-Improvement, 18, 251-258.

Bock, M., & Hormann, H. (1974). The influence of pictorial context upon retention of sentences: An investigation of Olson's cognitive theory of semantics. Psychologische Forschung, 36, 343-357.

Bock, M., & Milz,B. (1977). Pictorial context and the recall of pronoun sentences. Psychological Research, 39, 203-220.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. Cognitive Psychology, 11, 177-220.

Bower, G. H., Karlin, M. B., & Dueck, A. (1975). Memory for pictures. Memory and Cognition, 3, 216-220.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.

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

Britton, B. K., Glynn, S. M., Meyer, B. J. F., & Penland, M. J. (1982). Effects of text structure on use of cognitive capacity during reading. Journal of Educational Psychology, 74, 51-61.

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18, 32-42.

Carpenter, P. A., Miyake, A., & Just, M. A. (1995). Language comprehension: Sentence and discourse processing. Annual Review of Psychology, 46, 91-120.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8, 293-332.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., Glaser, R. (1989). Self-explanation: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). The nature of expertise. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. The Journal of the Learning Sciences, 1, 69-105.

Clements, P. (1976). The effects of staging on recall from prose. In R. Freedle (Ed.), Advances in discourse processing (Vol. 2) (pp. 287-330). Norwood, NJ: Ablex.

Collins, A. (1996). Design issues for learning environments. In S. Vosniadou, E. De Courte, R. Glaser, & H. Mandl (Eds.), International perspectives on the design of technology-supported learning environments, (pp. 347-361). Mahwah, NJ: Lawrence Erlbaum Associates.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. Discourse Processes, 25, 1-53.

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. Journal of Artificial Intelligence in Education, 6, 239-302.

Daneman, M., & Ellis, M. (1995). Memory for expository text: The effects of verbal and pictorial enhancements. Educational Psychology, 15, 115- 126.

Davidson-Shivers, G. V., Shorter, L., Jordan, K., & Ramussen, K. L. (1999). Learning strategies and navigation decisions of children using a hypermedia lesson. Jounal of Educational Multimedia and Hypermedia, 8, 175-188.

Deffner, G. (1988). Concurrent thinking aloud: An on-line tool for studying representations used in text understanding. Text, 8,351-367.

DeGroot, A. D. (1965). Thought and choice in chess. The Hague, The Netherlands: Mouton.

Denhière, G., & Denis, M. (1989). The processing of texts describing spatial configuration. In H. Mandl, & J. R. Levin, (Eds.), Knowledge acquisition from text and pictures (pp. 249-261). Amsterdam: Elsevier.

Deregowski, J. B. (1972). Reproduction of orientation of Kohs-type figures: A cross-cultural study. British Journal of Psychology, 63, 283-296.

Duchastel, P., & Waller, R. (1979). Pictorial illustration in instructional texts. Educational Technology, 19, 20-25.

Duchastel, P. (1978). Illustrating instructional texts. Educational Technology, 18, 36-39.

Dwyer, F. M. (1972). A guide for improving visualized instruction. State College, PA: Learning Services.

Dwyer, F. M. (1978). Strategies for improving visual learning: A handbook for the effective selection design, and use of visualized materials. State College, PA: Learning Services.

Dwyer, F. M. (1988). Examining the symbolic relationship between verbal and visual literacy in terms of facilitating student achievement. Reading Psychology: An international Quarterly, 9, 365-380.

Edstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for factor refrenced cognitive tests, Educational Testing Service, Princeton, NJ.

Ericsson, K. A. (1988). Concurrent verbal reports on text comprehension: A review. Text, 8, 235-295.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Ericsson, K. A., & Simon, H. A. (1985). Protocol analysis. In T. A. van Dijk (Ed.) Handbook of discourse analysis Volume 2: Dimensions of discourse (pp. 259-268). New York: Academic Press.

Ferguson-Hessler, M. G. M, & de Jong, T. (1990). Studying physic texts: Differences in study processes between good and poor performers. Cognition and Instruction, 7, 41-54.

Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. Journal of Memory and Language, 25, 43-58.

Fletcher, C. R. (1994). Levels of representation in memory for discourse. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics, (pp. 589-607). New York: Academic Press.

Fletcher, C. R., Chrysler, S. T. (1990). Surface forms, textbases, and situation models: Recognition memory for three types of textual information. Discourse Processes, 13, 175-190.

Franks, J. J., & Bransford, J. D. (1971). Abstraction of visual patterns. Journal of Experimental Psychology, 90, 65-74.

Frederiksen, C. H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 7, 371-458.

Frederiksen, C. H. (1986). Cognitive models and discourse analysis. In C. R. Cooper & S. Greenbaum (Eds.), Studying writing: Linguisitc approaches (pp. 227-268). Beverly Hills: Sage.

Frederiksen, C. H., Bédard, D., & Roy, M. C. (in preparation). An investigation of the integration of procedural and declarative knowledge for understanding and solving physics problems.

Frederiksen, C. H., Bracewell, R. J., Breuleux, A., & Renaud, A. (1989). The cognitive representation and processing of discourse: Function and dysfunction. In Y. Joanette & H. Brownell (Eds.), Discourse ability and brain damage: Theoretical and empirical perspectives (pp.). New York: Springer-Verlag.

Frederiksen, C. H., & Breuleux, A. (1990). Monitoring cognitive processing in semantically complex domains . In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), Diagnostic monitoring of skill and knowledge acquisition (pp. 351-391). Hillsdale, NJ: Erlbaum.

Frederiksen, C. H., & Donin, J. (1991). Constructing and manipulating representations in understanding and producing text. In G. Denhière and J. Rossi (Eds.), Text and text processing: Psychological and linguistic aspects (pp. 19-44). North Holland: Elsevier Science Publishers.

Frederiksen, C. H., & Donin, J. (1996, August). Conceptual processing in text comprehension and learning through discourse. Paper presented at an invited symposium at XXVI International Congress of Psychology, Montréal.

Garnham, A., Oakhill, J. V., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. Cognition, 11, 29-46.

Glenberg, A., & Langson, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. Journal of Memory and Language, 31, 129- 151.

Glenberg, A., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. Journal of Memory and Language, 26, 69-83.

Gobert, J. (1989). The interpretation of architectrual plans by expert and sub-expert architects. Unpublished Masters thesis, McGill University, Montreal.

Goldberg, F. (1974). Effects of imagery on learning incidental material in the classroom. Journal of Educational Psychology, 66, 233-237.

Goldstein, A. G., & Chance, J. E. (1971). Visual recognition memory for complex configurations. Perception & Psychophysics, 9, 237-241.

Goldstein, E. B. (1984). Sensation and perception: Second Edition, (pp. 29). Belmont, CA: Wadsworth Publishing Company.

Gombrich, E. (1990). Pictorial instructions. In H. Barlow, C. Blakemore, & M. Weston-Smith (Eds.), Images and understanding (pp. 26-45). New York, NY: Cambridge University Press.

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr, & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon, (pp. 285-318). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J.G. (1997). Understanding concepts in activity. In C. A. Weaver, S. Mannes, et al. (Eds.), Discourse comprehension: Essays in honor of Walter Kintsch, (pp, 65-95). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. American Psychologist, 53, 5-26.

Greenspan, S. L., & Segal, E. M. (1984). Reference and comprehension: A topic-comment analysis of sentence-picture verification. Cognitive Psychology, 16, 556-606.

Grimes, J. (1975). The thread of discourse. The Hague: Mouton.

Guri-Rozenblit, S. (1988). The interrelations between diagrammatic representations and verbal explanations in learning from social science texts. Instructional Science, 17, 219-234.

Haber, R. N., & Hershenson, M. (1973). The psychology of visual perception (pp. 19-27). Holt, Rinehart and Winston, Inc.

Haenggi, D., Kintsch, W., & Gernsbacher, M. S. (1995). Spatial situation models and text comprehension. Discourse Processes, 19, 173-199.

Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. Longman.

Haring, M. J., & Fry, M. A. (1979). Effect of pictures on children's comprehension of written texts. Educational Communication & Technology, 27, 185-190.

Hayes-Roth, B., & Thorndyke, B. W. (1979). Integration of knowledge from text. Journal of Verbal Learning and Verbal Behavior, 18, 91-108.

Hegarty, M. (1992). Mental animation: Motion from static displays of mechanical systems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 1084-1102.

Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl, & J. R. Levin, (Eds.), Knowledge acquisition from text and pictures (pp. 171-194). Amsterdam: Elsevier.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. Journal of Memory and Language, 32, 717-742.

Hegarty, M., Quilici, J., Narayanan, N. H., Holmquist, S., & Moreno, R. (1999). Multimedia instruction: Lessons from evaluation of a theory-based design. Jounral of Educational Multimedia and Hypermedia, 8, 119-150.

Holliday, W. G. (1975). The effects of verbal and adjunct pictorial-verbal information in science instruction. Journal of Research in Science Teaching, 12, 77-83.

Hong, E., & O'Neil Jr., H. F. (1992). Instructional strategies to help learners build relevant mental models in inferential statistics. Journal of Educational Psychology, 84, 150-159.

Hudson, W. (1960). Pictorial depth perception in sub-cultural groups in Africa. Journal of Social Psychology, 52, 183-208.

Huitema, J. S., Dopkins, S., Klin, C. M., & Myers, J. L. (1993). Connecting goals and actions during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 1053-1060.

Hutchins, E. (1980). Cultrure and inference. Cambridge, MA: Harvard University Press.

Irwin, J. W. (1980). The effect of linguistic cohesion on prose comprehension. Journal of Reading Behavior, 12, 325-332.

Johnson-Laird, P. N. (1983). Mental models. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jörg, S., & Hörmann, H. (1978). The influence of general and specific verbal labels on the recognition of labveled and unlabeled parts of pictures. Journal of Verbal Learning and Verbal Behavior, 17, 445-4545.

Just, M. A.,& Carpenter, P. A. (1987). The psychology of reading and language comprehension. Boston: Allyn and Bacon.

Keeton, W. T. (1976). Biological science (Third Edition), (pp. 441). New York, NY: W. W. Norton & Company.

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

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

Kintsch, W. (1989). Learning from text. In L. B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 25-46). Hillsdale, NJ: Erlbaum.

Kintsch, W. (1994). Text comprehension, memory, and learning. American Psychologist, 49, 294-303.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. New York, NY: Cambridge University Press.

Kintsch, W., Kozminsky, E., Streby, W., McKoon, G., & Keenan, J. M. (1975). Comprehension and recall as a function of content variables. Journal of Verbal Learning and Verbal Behavior, 14, 196-214.

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

Kosslyn, S. M., & Pomerantz, J. R. (1977). Imagery, propositions, and the form of internal representations. Cognitive Psychology, 9, 52-76.

Kozma, R. B. (1991). Learning with media. Review of Educational Research, 61, 179-211.

Kubes, M. (1988). Use of prior knowledge in integration of information from technical materials. Unpublished doctoral dissertation, McGill University, Montreal.

Kuntz, G. C., Drewniak, U., & Schott, F. (1992). On-line and off-line assessment of self-regulation in learning from instructional text. Learning and Instruction, 2, 287-301.

Larkin, J. H. (1989). Display-based problem solving. In D. Klahr & K. Kotovsky (Eds.). Complex information processing: The impact of Herbert A. Simon, (pp. 319-341). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-99.

László, Meutsch & Viehoff, (1988). Verbal reports as data in text comprehension research: An introduction. Text, 8, 283-294.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), The nature of expertise (pp. 311-342). Hillsdale, NJ: Lawrence Erlbaum Associates.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. Educational Communication and Technology: A Journal of Theory, Research, and Development, 30, 195-232.

Levin, J. R. (1981a). The mnemonic '80s: Keywords in the classroom. Educational Psychologist, 16, 65-82.

Levin, J. R. (1981b). On the functions of pictures in prose. In F. J. Pirozzolo & M. C. Wittrock (Eds.), Neuropsychological and cognitive processes in reading (pp. 203-228). New York: Academic Press.

Levin, J. R. (1989). A transfer-appropriate-processing perspective of pictures in prose. In H. Mandl & J. R. Levin (Eds.), Knowledge acquisition from text and pictures: Advances in psychology, 58, (pp. 83-100). Amsterdam: Elsevier.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. A. Willows & H. A. Houghton (Eds.), The psychology of illustration: Basic Research (Vol. 1) (pp. 51-85). London: Springer-Verlag.

Levin, J. R., Bender, B. G., & Lesgold, A. M. (1976). Pictures, repetition, and young children's oral prose learning. AV Communication Review, 4, 367-380.

Levin, J. R., Bender, B. G., & Pressley, (1979). Pictures, imagery, and children's recall of central versus peripheral sentence information. Educational Communcation and Technology, 27, 89-95.

Levin, J. R., & Lesgold, A. (1978). On pictures in prose. Educational Communications and Technology, 26, 233-243.

Long, D. L., & Bourg, T. (1996). Thinking-aloud: Telling a story about a story. Discourse Processes, 21, 329-339.

Lowe, R. K. (1993). Constructing a mental representation from an abstract technical diagram. Learning and Instruction, 3, 157-179.

Mandler, J. M., & Johnson, N. S. (1976). Some of the thousand words a picture is worth. Journal of Experimental Psychology: Human Learning and Memory, 2, 529-540.

Mandler, J. M., & Parker, R. E. (1976). Memory for descriptive and spatial information in complex pictures. Journal of Experimental Psychology: Human Learning and Memory, 2, 38-48.

Mannes, S. (1994). Strategic processing of text. Journal of Educational Psychology, 86, 577-588.

Mannes, S. M., & Kintsch, W. (1987). Knowledge organization and text organization. Cognition and Instruction, 4, 91-115.

Marshall, N., & Glock, M. D. (1978). Comprehension of connected discourse: A study into the relationship between the structure of test and information recalled. Reading Research Quarterly, 14, 10-56.

Mayer, R. E. (1993a). Illustrations that instruct. In R. Glasser (Ed.), Advances in instructional psychology (Vol. 4) (pp. 253-284). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mayer, R. E. (1993b). Commentary. Comprehension of graphics in texts: an overview. Learning and Instruction, 3, 239-245.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? Educational Psychologist, 32, 1-19.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. Journal of Educational Psychology, 83, 484-490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. Journal of Educational Psychology, 84, 444-452.

Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraing and signaling. Journal of Educational Psychology, 76, 1089-1105.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? Journal of Educational Psychology, 82, 715-726.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. Journal of Educational Psychology, 90, 312-320.

Mayer, R. E., & Sims, V. (1993, April). Role of spatial ability in learning with computer-generated animations. Paper presented at the Annual Meeting of the American Educational Research Association. Atlanta.

McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. Reading Research Quarterly, 27, 79-93.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. Psychological Review, 99, 440-466.

McNamara, T. P. (1986). Mental representations of spatial relations. Cognitive Psychology, 18, 87-121.

McNamara, D. S., Kintsch, E., Butler Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. Cognition and Instruction, 14, 1-43.

Moe, A. J. (1979). Cohesion, coherence, and the comprehension of text. Journal of Reading, 23, 16-20.

Moore, P. J. (1993). Metacognitive processing of diagrams, maps and graphs. Learning and Instruction, 3, 215-226.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative compehension. Journal of Memory and Language, 26, 165-187.

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 777-784.

Olson, G. M., Duffy, S. A., & Mack,R. L. (1984). Thinking-out-loud as a method for studying real-time comprehension processes. In D. Kieras & M. Just (Eds.), New methods in the study of immediate processes in comprehension (pp. 253-286). Hillsdale, NJ: Erlbaum.

Oshima, J., Scardamalia, M., & Bereiter, C. (1996). Collaborative learning processes associated with high and low conceptual progress. Instructional Science, 24, 125-155.

Paivio, A. (1986). Mental representations: A dual coding approach. Oxford, England: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45, 255-287.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. Cognition and Instruction, 1, 117-175.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. Memory & Cognition, 3, 519-526.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp. 259-303). Hillsdale, NJ: Erlbaum.

Peeck, J. (1980, April). Experimenter-provided and learner-generated pictures in learning from text. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Peeck, J. (1987). The role of illustrations in processing and remembering illustrated text. In D. M. Willows, & H. A. Houghton, (Eds.), The psychology of illustration: Basic research (Vol. 1.) (pp. 113-151). New York, NY: Springer-Verlag.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. Learning and Instruction, 3, 227-238.

Perfetti, C. A. (1985). Reading ability. New York, NY: Oxford University Press.

Petre, M., & Green, T. (1992). Requirements of graphical notations for professional readers. Le Travail Humain, 55, 47-70.

Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. Journal of Memory and Language, 24, 503-518.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), Artificial intelligence and the future of testing (pp. 73-126), Hillsdale, NJ: Erlbaum.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. Journal of Experimental Psychology: General, 115, 281-294.

Potts, G. R. (1977). Integrating new and old information. Journal of Verbal Learning & Verbal Behavior, 16, 305-320.

Potts, G. R., St. John, M., & Kirson, D. (1989). Incorporating new information into existing world knowledge. Cognitive Psychology, 21, 303-333.

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

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 80, 1-24.

Rankin, E. F., & Culhane, J. N. (1970). One picture equals 1,000 words? Reading Improvement, 7, 37-40.

Rasco, R. W., Tennyson, R. D., & Boutwell, R. C. (1975). Imagery instructions and drawings in learning prose. Journal of Educational Psychology, 67, 188-192.

Readence, J. R., & Moore, D. W. (1981). A meta-analytic review of the effect of adjunct pictures on reading comprehension. Psychology in the Schools, 18, 218-224.

Reid, D., & Beverridge, M. (1986). Effects of text illustration on children's learning of a school science topic. British Journal of Educational Psychology, 56, 294-303.

Reid, D., & Beverridge, M. (1990). Reading illustrated science texts: A micro-computer based investigation of children's strategies. British Journal of Educational Psychology, 60, 76-87.

Reid, D., Briggs, N., & Beverridge, M. (1983). The effect of picture upon the readability of a school science topic. British Journal of Educational Psychology, 53, 327-335.

Reinking, D., Hayes, D. A., & McEneaney, J. E. (1988). Good and poor readers' use of explicityly cued graphic aids. Journal of Reading Behavior, 20, 229-247.

Renaud, A. & Frederiksen, C. H. (1988). On-line processing of a procedural text. In Proceedings of The Tenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.

Resnick, L. B. (1987). Learning in school and out. Educational Researcher, 16, 13-20.

Rieber, L. P. (1990a). Using computer animated graphics in science instruction with children. Journal of Educational Psychology, 82, 135-140.

Rieber, L. P. (1990b). Animation in computer-based instruction. Educational Technology, Research & Development, 38, 77-86.

Rieber, L. P. (1991a). Animation, incidental learning, and continuing motivation. Journal of Educational Psychology, 83, 318-328.

Rieber, L. P. (1991b). Effects of visual grouping strategies of computer-animated presentations on selective attention in science. Educational Technology, Research & Development, 39, 5-15.

Rinck, M., Williams, P., Bower, G. H., & Becker, E. S. (1996). Spatial situation models and narrative understanding: Some generalizations and extensions. Discourse Processes, 21, 23-55.

Roy, M. C. (1991). The effects of prior knowledge and staging on the processing and comprehension of procedural text. Unpublished Masters thesis, McGill University, Montreal.

Royer, J. M., & Cable, G. W. (1976). Illustrations, analogies, and facilitative transfer in prose learning. Journal of Educational Psychology, 68, 205-209.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. Laberge & J. Samuels (Eds.), Basic processes in reading: Perception and comprehension (pp. 265-303). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, and W. F. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rusted, J. M., & Coltheart, V. (1979). Facilitation of children's prose recall by the presence of pictures. Memory and Cognition, 7, 354-359.

Rusted, J., & Hodgson, S. (1985). Evaluating the picture facilitation effect in children's recall of written texts. British Journal of Educational Psychology, 55, 288-294.

Salomon, G. (1989). Learning from texts and pictures: Reflections on a meta-level. In H. Mandl, & J. R. Levin, (Eds.), Knowledge acquisition from text and pictures (pp. 73-82). Amsterdam: Elsevier.

Samuals, S. J. (1967). Attentional process in reading: The effect of pictures on the acquisition of reading responses. Journal of Educational Psychology, 58, 337-342.

Samuals, S. J. (1970). Effects of pictures on learning to read, comprehension and attitude. Review of Educational Research, 40, 398-407.

Schallert, D. L. (1980). The role of illustrations in reading comprehension. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in prose comprehension, (pp. 503-524). Hillsdale, NJ: Lawrence Earlbaum Associates.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding. Hillsdale, NJ: Lawrence Earlbaum Associates.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. Journal of Memory and Language, 25, 279-294.

Schnotz, W. (1993a). Introduction. Learning and Instruction, 3, 151-155.

Schnotz, W. (1993b). Some remarks on the commentary. On the relation of dual coding and mental models in graphics comprehension. Learning and Instruction, 3, 247-249.

Schnotz, W., Picard, E., & Hron, A. (1993). How do successful and unsuccessful learners use texts and graphics? Learning and Instruction, 3, 181-199.

Sherman, J. L. (1976). Contextual information and prose comprehension. Journal of Reading Behavior, 8, 369-379.

Shriberg, L. K., Levin, J. R., McCormick, C. B., & Pressley, M. (1982). Learning about "famous" people via the keyword method. Journal of Educational Psychology, 74, 238-247.

Small, M. Y., Dowling, S., & Lovett, S. B. (1981, April). Semantic elaboration and children's recal of picture and prose information. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Small, M. Y., Lovett, S. B., & Scher, M. S. (1993). Pictures facilitate children's recall of unillustrated expository prose. Journal of Educational Psychology, 85, 520-528.

Stenning, K., & Oberlander, J. (1994). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. Cogitive Science, 19, 97-140.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. Cognitive Psychology, 10, 422-437.

Stone, D. E., & Glock, M. D. (1981). How do young adults read directions with and without pictures? Journal of Educational Psychology, 73, 419-426.

Suchman, L. (1988). Plans and situated actions: The problem of human/machine communication. Cambridge, UK: Cambridge University Press.

Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. Journal of Memory and Language, 32, 279-300.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. Cognition and Instruction, 8, 351-362.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. Journal of Experimental Psychology: General, 119, 176-192.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. Journal of Memory and Language, 31, 261-292.

Thorndyke, P. W., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. Cognitive Psychology, 12, 137-175.

Tversky, B. (1981). Distortions in memory for maps. Cognitive Psychology, 13, 407-433.

Twyman, M. (1985). Using pictorial language: A discussion of the dimensions of the problem. In T. M. Duffy, & R. Waller (Eds.), Designing usable texts (pp. 245-312). New York, NY: Academic Press.

van Dijk, T. A. (1985). Handbook of discourse analysis Vol. 2: Dimensions of discourse. San Diego, CA: Academic Press.

van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. San Diego, CA: Academic Press.

Vernon, M. D. (1953). The value of pictorial illustration. British Journal of Educational Psychology, 23, 180-187.

Waddill, P. J., McDaniel, M. A., & Einstein, G. O. (1988). Illustrations as adjuncts to prose: A text-appropriate processing approach. Journal of Educational Psychology, 80, 457-464.

Walker, C. H., & Meyer, B. J. F. (1980a). Integrating information from text: An evaluation of current theories. Review of Educational Research, 50, 421-437.

Walker, C. H., & Meyer, B. J. F. (1980b). Integrating different types of information in text. Journal of Verbal Learning and Verbal Behavior, 19, 263-275.

Weidenmann, B. (1989). When good pictures fail: An information-processing approach to the effect of illustrations. In H. Mandl, & J. R. Levin, (Eds.), Knowledge acquisition from text and pictures (pp. 157-170). Amsterdam: Elsevier.

Weintraub, S. (1960). The effect of pictures on the comprehension of a second grade basal reader. Unpublished doctoral dissertation, University of Illinois.

Weisberg, J. S. (1970). The use of visual advance organizers for learning earth science concepts. Journal of Research in Science Teaching, 7, 161-165.

Whitney, P., & Budd, D. (1996). Think-aloud protocols and the study of comprehension. Discourse Processes, 21(3), 341-351.

Williams, C. M. (1968). Learning from pictures. Washington, D.C.: Association for Educational Communications and Technology.

Willows, D. M. (1978). A picture is not always worth a thousand words: Pictures as distractors in reading. Journal of Educational Psychology, 70, 255-262.

Wilson, S. G., Rinick, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1993). Mental models and narrative comprehension: Some qualifications. Journal of Memory & Language, 32, 141-154.

Winn, W. (1987). Charts, graphs, and diagrams in educational materials. The psychology of illustration: Basic research (Vol. 1) (pp. 152-198). London: Springer-Verlag.

Yee, P. L., Hunt, E., & Pellegrino, J. W. (1991). Coordinating cognitive information: Task effects and individual differences in integrating information from several sources. Cognitive Psychology, 23, 615-680.

Zhang, J. (1997). The nature of external representations in problem solving. Cognitive Science, 21, 179-217.

Zwaan, R. A., & Van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension? Discourse Processes, 16, 125-143.

Appendix A

Pre-Experimental Questionnaire

Pretask Questionnaire

Please answer the following questions to the best of your ability. The purpose of this brief questionnaire is to determine your level of familiarity with the topic that you will be reading about.

1) Have you ever taken a course(s) that dealt with the topic of the human visual system?

2) If you have answered yes to question one (above), please indicate the approximate course title(s), the grade level of the course (e.g., grade 10 biology, undergraduate psychology, etc.), how long ago you took the course(s), and approximately how many classes were devoted to the topic.

<u>Title ______________ Level ____________________ Date ______________ # Classes</u>

3) In a brief paragraph, please describe your current understanding of the human visual system.

Appendix B

Post-Experimental Questionnaire

Questionnaire

Date: ______________________ Time:____________

Program:___________________ Year:____________

1) Generally speaking, how familiar were you with the content of this text?

1	2	3	4	5	6	7
totally unfamiliar						very familiar

2) What concepts (if any) were you already familiar with?

3) Generally speaking, how hard or easy to understand was the text ?

1	2	3	4	5	6	7
very hard						very easy

4) How helpful do you think the illustrations were/would be in helping you to understand this text?

1	2	3	4	5	6	7
not at all helpful						extremely helpful

5) Briefly describe how they helped/would help you.

6) How do you think you could have improved your comprehension of the text?

7) Do you think the diagrams presented in this study could have been improved to help you comprehend the text? If so, how?

8) How often do you encounter similar types of diagrams (i.e., diagrams depicting the parts of a functional system) in your area of study?

1	2	3	4	5	6	7
never						all the time

9) Estimate the percentage of space that you think each of the following materials in your area of study devote to such diagrams?

	Estimated %
Textbooks	____________
Journal Articles	____________
Manuals	____________
Other(s) ____________________	____________

10) When you encounter such diagrams in your area of study how often do you use them to help you understand the text?

1	2	3	4	5	6	7
never						all the time

11) Generally speaking, how hard or easy do you find such diagrams to understand?

1	2	3	4	5	6	7
very hard						very easy

Appendix C

Correlations Between Spatial Ability and Dependent Measures

Correlations between Visual Recall and Spatial Ability

	Structure	Function	Energy
SA_T1	0.307	-0.104	-0.080
SA_T2	0.100	0.231	-0.248

Correlationss between Mental Model Nodes and Spatial Ability

	Structure	Function	Energy
SA_T1	0.363	0.265	-0.048
SA_T2	0.211	0.346	0.258

Correlations between Mental Model Relations and Spatial Ability

	Structure	Function	Energy
SA_T1	0.273	0.274	0.057
SA_T2	0.187	0.222	0.236

Correlations between Verbal Recall and Spatial Ability

	Structure	Function	Energy
SA_T1	0.243	0.149	0.182
SA_T2	0.357	0.314	0.143

Correlations Between Proportion of Acitive On-line Processes and Spatial Ability

	On-Line
SA_T1	0.013
SA_T2	0.211

Correlations between Reading Time and Spatial Ability

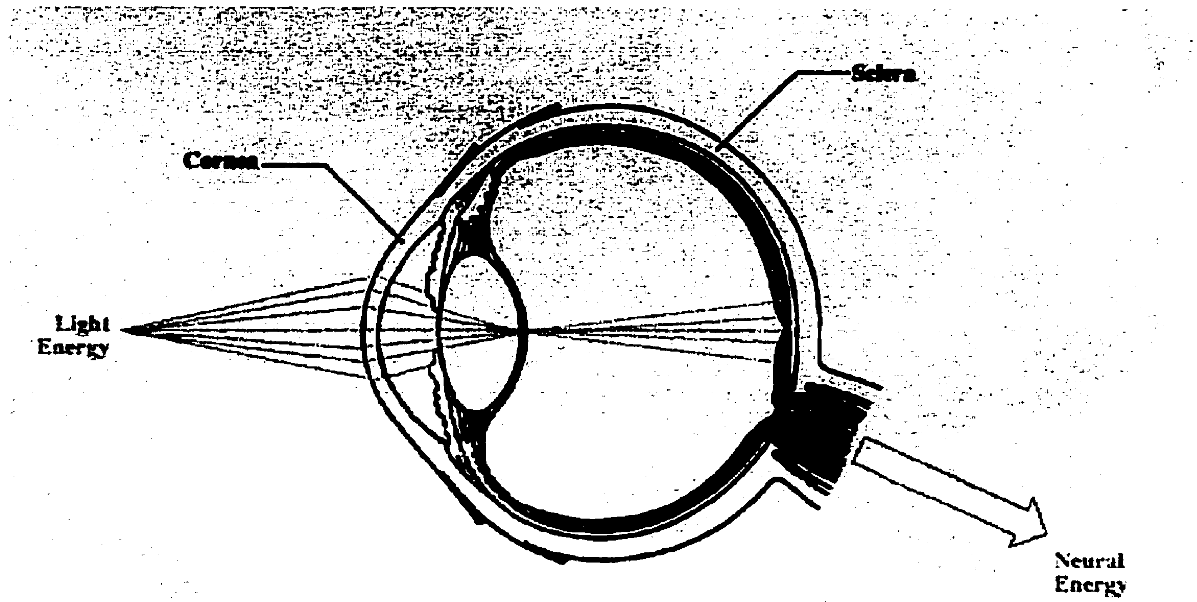
	Structure	Function	Energy
SA_T1	-0.016	-0.086	-0.140
SA_T2	0.104	-0.064	0.309

Note. * = $\alpha < .05$; ** = $\alpha < .01$; SA_T1 = Paper Folding Test; SA_T2 = Surface Development Test.

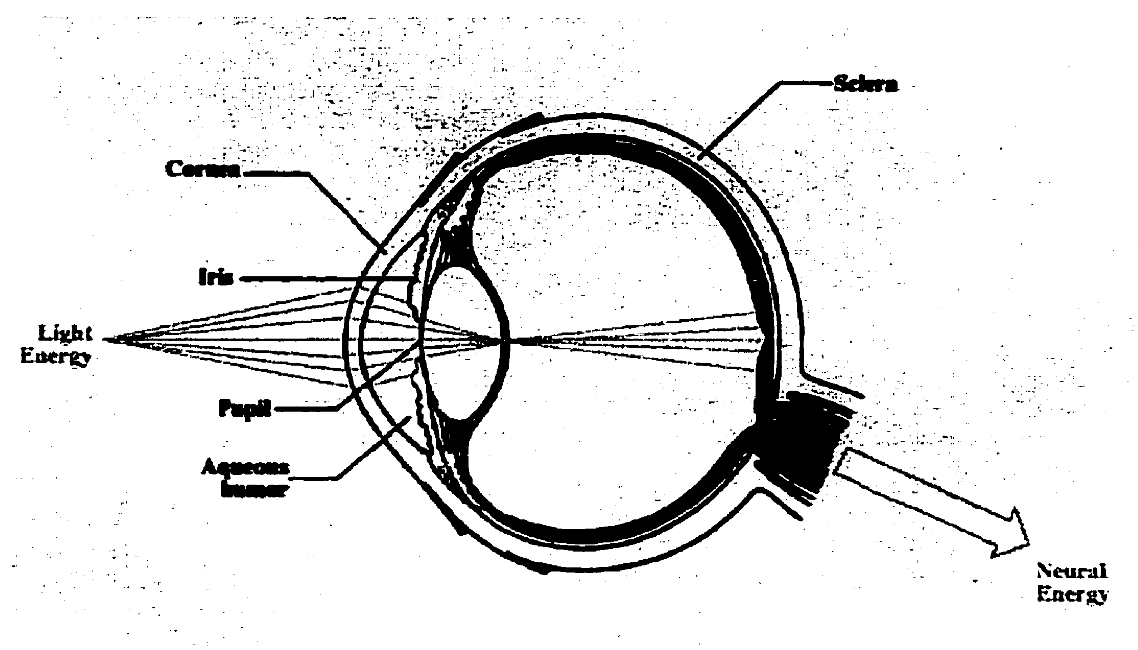
Appendix D

Stimulus Illustrations

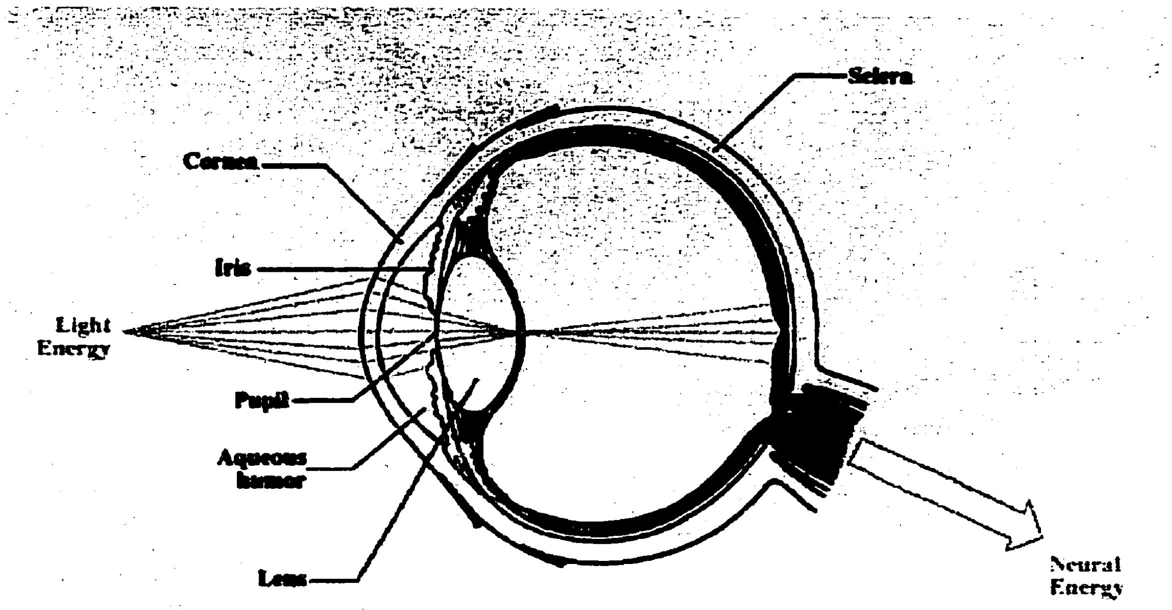
(a) Cross section of the eye: Illustration 1.1



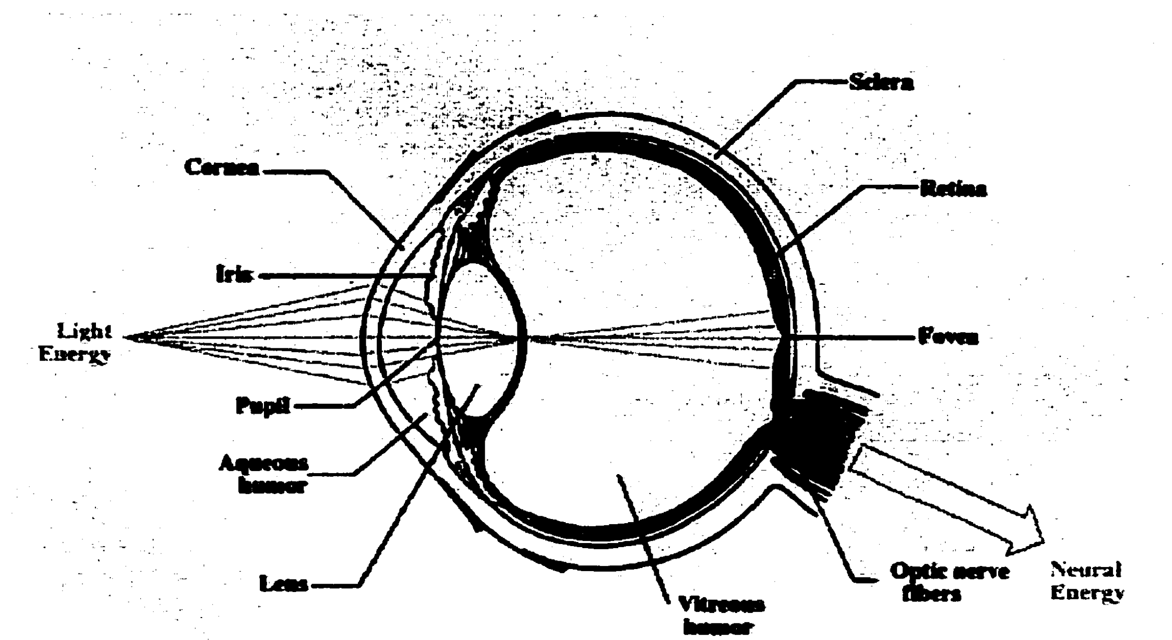
(b) Cross section of the eye: Illustration 1.2



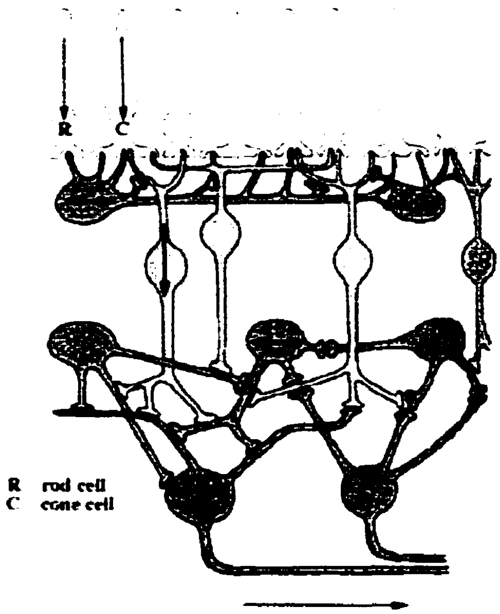
(c) Cross section of the eye: Illustration 1.3



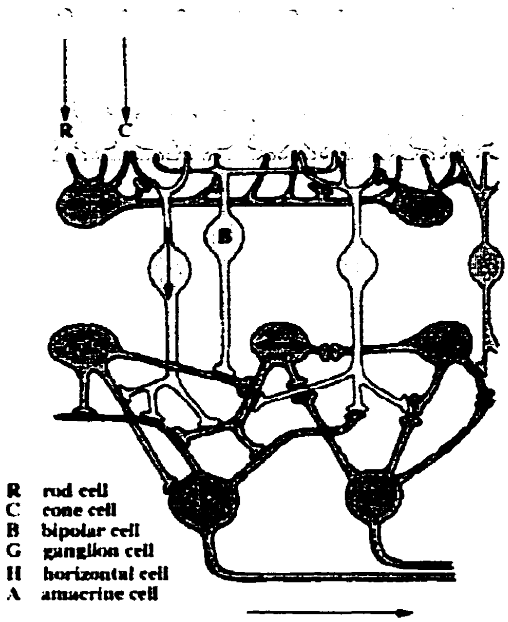
(d) Cross section of the eye: Illustration 1.4



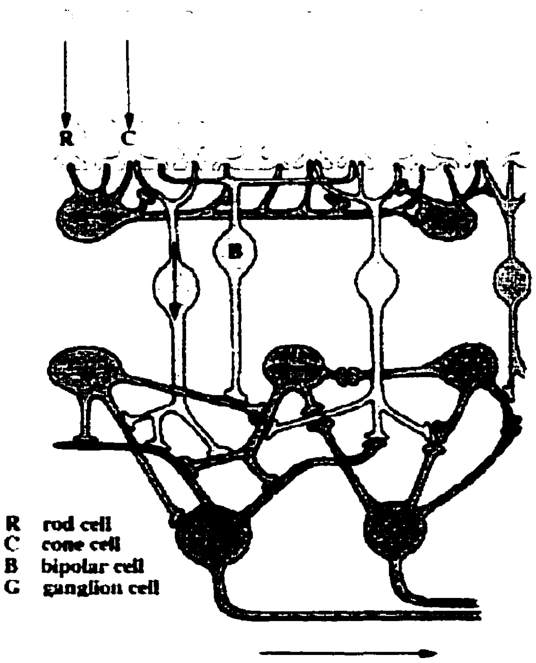
(e) Enlarged view of retinal cells: Illustration 2.1



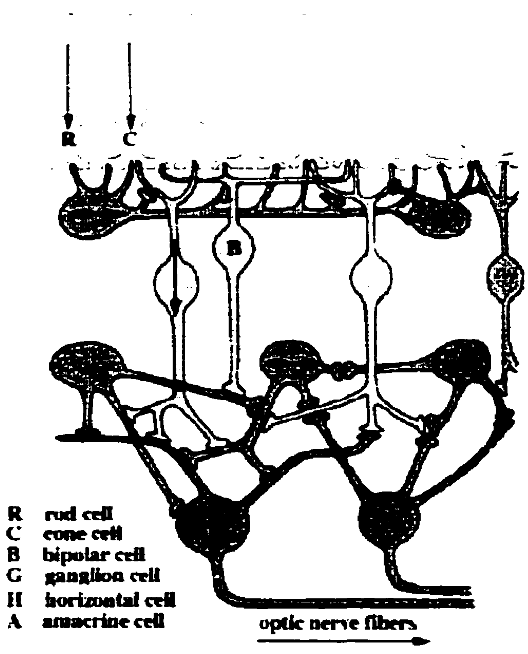
(g) Enlarged view of retinal cells: Illustration 2.3



(f) Enlarged view of retinal cells: Illustration 2.2



(h) Enlarged view of retinal cells: Illustration 2.4



Appendix E

Experimental Text

The Human Visual System

+Paragraph 1

We will discuss the principal components of the human eye. [structure]

The eyeball is about 25 millimeters in diameter. [structure]

The opaque protective layer that surrounds the eyeball is called the sclera. [structure]

The outside opening at the front of the eye is covered by a clear membrane called the cornea. [structure]

/picture 11

The shape of the cornea is responsible for about 70 percent of the eye's focusing power. [function]

Light first passes through this structure on its way to the retina. [energy]

The cornea must be transparent and free from scar tissue that may result from injury, infection, or allergic reactions in order for a sharp image to be formed on the retina. [other]

+Paragraph 2

Behind the cornea is the iris. [structure]

The iris is a sphincter muscle which can open and close, and is thus able to adjust the amount of light that can enter the eye. [function]

The iris is the pigmented part of the front of the eye. [structure]

The pupil is an opening located in the center of the iris. [structure]

/picture 12

After light exits from the cornea it next passes through the pupil. [energy]

The iris contains smooth muscle fibers arranged in both circular and radial directions. [structure]

The pupil is reduced when the circular muscles contract. [function]

The pupil is dilated when the radial muscles contract. [function]

The iris thus regulates the size of the pupil, admitting light in about the same way as a diaphragm of a camera regulates the lens aperture. [function]

+Paragraph 3

Behind the cornea is the lens. [structure]

The lens consists of a number of transparent layers arranged much like the layers of an onion. [structure]

These layers are held in place in a sac. [structure]

/picture 13

After light passes through the pupil, it next encounters the lens. [energy]

Although most of the refraction or bending of light is done by the cornea, some remaining focusing is accomplished by varying the thickness and hence the refractive power of the lens. [function]

This process is called accommodation. [other]

In this process, the ciliary muscles tighten and thereby increase the focusing power of the eye. [function]

The curvature of the lens is increased by this tightening of muscles. [function]

Accommodation enables us to keep an image on the retina sharp as we look at objects located at different distances. [function]

+Paragraph 4

The vitreous humor, a clear jellylike substance, is located in the middle of the eyeball. [structure]

The purpose of this substance is to maintain the shape of the eye, and also to provide a medium with a similar refractive index to that of the lens, so that no further refraction of light occurs. [function]

After the light exits from the lens it passes through the vitreous humor. [energy]

The light is then focused on the inside surface of the ball. [energy]

The retina is an area of the inside surface of the eyeball that surrounds nearly 200 degrees as measured on the circumference of a circle. [structure]

The photoreceptors and their neural support are embedded in the retina. [structure]

The fovea is an area of the retina located directly behind the lens. [structure]

/picture 14

It is the most sensitive portion of the retina to detecting light patterns. [function]

The blind spot, or optic disk, is an area on the retina that covers several degrees. [structure]

It has no receptors because the nerve connections from the receptors exit from the eye at this point to form the optic nerve tract connecting to higher centers in the brain. [structure]

+Paragraph 5

The photoreceptors convert light energy into neural energy. [function]

They are embedded in the back of the retina, with all of the neural connections and blood supply in front of them. [structure]

/picture 21

Consequently, the light passes through all of the supporting structures before reaching the receptors. [energy]

There are two kinds of photoreceptors embedded in the back of the retina called rods and cones. [structure]

Rods are the rod-shaped receptors that are primarily responsible for vision in low levels of illumination. [function]

The rod system is extremely sensitive in the dark, but cannot resolve fine details. [function]

Cones are cone-shaped receptors that are primarily responsible for vision in high levels of illumination. [function]

The cone system is responsible for detail vision and color vision. [function]

+Paragraph 6

There are about 120 million rods and nearly seven million cones. [structure]

Most of the cones are contained in the center of the fovea, which covers an area of perhaps 1 degree. [structure]

The number and proportion of cones falls off rapidly with many fewer cones present beyond 10 degrees. [structure]

On the other hand, there are no rods in the center of the fovea. [structure]

They reach their highest frequency at about 16 degrees on either side, with decreasing numbers out to about 100 degrees at the edge of the retina on either side of the fovea. [structure]

+Paragraph 7

The photoreceptors synapse onto bipolar and then ganglion cells. [structure]

/picture 22

Electrical signals start in the photoreceptors and travel to the brain via these synaptic connections. [energy]

In the fovea, usually only one cone is connected to one bipolar cell. [structure]

Outside of the fovea there will be many photoreceptors connected to one bipolar cell. [structure]

The number of photoreceptors converging onto a single bipolar cell approaches a convergence of hundreds of rods onto one bipolar beyond 20 degrees into the periphery. [structure]

This great pooling of receptors onto a single bipolar cell in the periphery means that any particular bipolar cell cannot determine which of its many receptors has been stimulated by light. [structure & function]

On the other hand, in the center of the fovea there is very little pooling, resulting in virtually perfect specificity of excitation. [function]

+Paragraph 8

Each bipolar cell is connected to a ganglion cell via a second synapse. [structure]

The ganglion cell has an elongated body that forms one of the fibers of the optic nerve. [structure]

There are less than one million ganglion cells leaving the retina in this manner. [structure]

Again, in the fovea each bipolar cell generally connects to one ganglion, whereas in the periphery a number of bipolars will converge on a single ganglion cell. [structure]

The two remaining types of retinal cells are horizontal and amacrine cells. [structure]

Horizontal cells connect receptors to other receptors. [structure]

Amacrine cells connect ganglion cells to other ganglion cells and also bipolar cells to other bipolar cells. [structure]

/picture 23

These cells do not transmit signals towards the brain, but instead they transmit and pool signals laterally across the retina. [function]

Thus, neural signals flow both directly towards the brain and laterally across the retina before going to the brain. [energy]

+Paragraph 9

We still do not completely understand the details concerning how the energy in photons is transduced into electrical activity in neurons. [other]

The first stage requires absorption of a photon by a molecule of photopigment in a photoreceptor. [function]

The photopigment in rod receptors is called rhodopsin, or visual purple. [structure]

/picture 24

Each of the four million or so molecules of rhodopsin in each rod will undergo a molecular change upon the absorption of a photon. [function]

This change occurs almost instantaneously with absorption. [function]

The products of this change are capable of producing a neural charge. [function]

This charge flows across the synapse between the photoreceptor and the bipolar cell it connects to. [energy]

Note. / = locations of forced exposure to an illustration for participants in the Controlled Access condition; + = locations where a new paragraph begins; [] = information type for each sentence;

Appendix F

Propositional Analysis of Text

The Human Visual System

+Paragraph 1

We will discuss the principal components of the human eye. [structure]

S	2	ACT:Discuss AGT:we TNS:future THM: 2.1,2.2
S	2.1	OBJ:component STATE.ID.PRT:eye ATT:principal NUM:plural
S	2.2	OBJ:eye STATE.ID.ATT:huma n

The eyeball is about 25 millimeters in diameter. [structure]

S	3	OBJ:eyeball ATT:diameter DGR:about MEAS:25 UNIT:millimeter

The opaque protective layer that surrounds the eyeball is called the sclera. [structure]

S	4	PRC:surround PAT:4.1 REL.OBJ:eyeball
S	4.1	OBJ:layer ATT:opaque ATT:protective DET:generic NUM:singular
S	4.2	EQUIV:"called" [4.1] [sclera]

The outside opening at the front of the eye is covered by a clear membrane called the cornea. [structure]

S	5	OBJ:eye ST.ID.PRT:5.1,5.2
S	5.1	OBJ:opening

		LOC:outside
		DIR:front
S	5.2	PRC:cover
		PAT:opening
		INST:5.3
S	5.3	OBJ:membrane
		ATT:clear
		DET:definite
		NUM:singular
S	5.4	EQUIV:[cornea]
		[5.3]
		called

/picture 11

The shape of the cornea is responsible for about 70 percent of the eye's focusing power. [function]

F	6	CAU:*responsible*
		[6.1]
		[6.2]
F	6.1	OBJ:cornea
		ATT:shape
F	6.2	ACT:focus
		AGT:eye
		ACT.ID.ATT:power
		DEG:about
		MEAS:70
		UNIT:percent

Light first passes through this structure on its way to the retina. [energy]

E	7	ACT:pass
		AGT:light
		R.OBJ:7.1
		SRC:7.2
		RSLT:7.3
		DIR:through
E	7.1	OBJ:structure
		DEF:this
		NUM:singular
E	7.2	OBJ:light
		LOC:cornea
E	7.3	OBJ:light
		LOC:retina
E	7.4	ORD:LOC:
		[7.2]
		[7.3]

The cornea must be transparent and free from scar tissue that may result from injury, infection, or allergic reactions in order for a sharp image to be formed on the retina. [other]

	8	COND:[8.1, 8.2] [8.6, 8.10] "in order for"
	8.1	ACT:form AGT:empty OBJ:image RSLT:8.2
	8.2	OBJ:image ATT:sharp REL.OBJ:retina DET:definite NUM:singular LOC:on
	8.3	OBJ:cornea ATT:transparent MOD:must
	8.4	POSS:[cornea] [8.5] TRTH:neg
	8.5	OBJ:tissue ATT:scar]
	8.6	CAU:[8.5] [injury]
	8.7	CAU: [8.5] [infection]
	8.8	CAU: [8.5] [8.9]
	8.9	ACT: reaction ACT.ATT:allergic
	8.10.	AND:[8.6] [8.7] [8.8]

+Paragraph 2

Behind the cornea is the iris. [structure]

S	9	OBJ:iris LOC:behind R.OBJ:cornea

The iris is a sphincter muscle which can open and close, and is thus able to adjust the amount of light that can enter the eye. [function]

F	**10.0.**	PRC:adjust PAT:iris R.OBJ:light R.ACT
F	**10.1**	ACT:enter AGT:light AGT.ATT:amount R.OBJ:eye MOD:can
F	**10.2**	OBJ:light ATT:amount
F	**10.3**	EQUIV:[iris] [muscle]
F	**10.4**	OBJ:muscle STATE.ID.ATT:sphint er
F	**10.5**	ACT:open AGT:iris MOD:can RSLT:10.0
F	**10.6**	ACT:close AGT:iris RSLT:10.0
F	**10.7**	AND:[10.5] [10.6]

The iris is the pigmented part of the front of the eye. [structure]

S	**11.0.**	PRT:[iris] [eye]
S	**11.1**	OBJ:iris ATT:pigmented R.OBJ:eye DET:definite NUM:singular LOC:front
S	**11.2**	EQUIV:[iris] [11.0, 11.1]

The pupil is an opening located in the center of the iris. [structure]

S	**12**	OBJ:iris STATE.ID.PRT:openin g LOC:center

S	12.1	EQUIV:"is" [pupil] [12.0]

/picture 12

After light exits from the cornea it next passes through the pupil. [energy]

E	13	ACT:exit AGT:light SRC:13.1 DIR:"from"
E	13.1	OBJ:light LOC:cornea
E	13.2	ACT:passes AGT:it (light) OBJ:pupil RSLT:13.3 DIR:"through"
E	13.3	OBJ:light LOC:pupil TRTH:neg
E	13.4	ORD.LOC:[13.0] [13.2] "after" "next"

The iris contains smooth muscle fibers arranged in both circular and radial directions. [structure]

S	14	PRT:[iris] [14.1]
S	14.1	PRC:contain PAT:iris DET:definite NUM:singular OBJ:14.2
S	14.2	OBJ:fiber DET:definite NUM:plural STATE.ID.ATT:smoot h muscle
S	14.3	ACT:arrange AGT:empty OBJ:14.2 ATT:direction DIR:circular
S	14.4	ACT:arrange

		AGT:empty
		OBJ:14.2
		ATT:direction
		DIR:radial
S	14.5	AND
		[14.3]
		[14.4]

The pupil is reduced when the circular muscles contract. [function]

F	15	COND:[15.1]
		[15.3]
F	15.1	ACT:contract
		AGT:15.2
F	15.2	OBJ:muscle
		ATT:cicular
		DET:definite
		NUM:plural
F	15.3	ACT:reduce
		OBJ:pupil
		DET:definite
		NUM:singular
		VCE:passive

The pupil is dilated when the radial muscles contract. [function]

F	16	COND:[16.1]
		[16.3]
F	16.1	ACT:contract
		AGT:16.2
F	16.2	OBJ:muscle
		ATT:radial
		DET:definite
		NUM:plural
F	16.3	ACT:dilate
		OBJ:pupil
		DET:definite
		NUM:singular
		VCE:passive

The iris thus regulates the size of the pupil, admitting light in about the same way as a diaphragm of a camera regulates the lens aperture. [function]

F	17	COND:[12.0,
		13.0]
		[17.1]
F	17.1	PRC:regulate
		PAT:iris
		DET:definite
		NUM:singular

		R.OBJ:17.2 REL.ACT:14.3
F	17.2	OBJ:pupil DET:definite NUM:singular ATT:size
F	17.3	ACT:admit AGT:elided(iris) R.OBJ:light ASP:continuing PROX.ACT.ATT:*in about the same way as* 17.4
F	17.4	PRC:regulate PAT:diaphram DET:definite NUM:singular R.OBJ 17.5
F	17.5	OBJ:aperture STATE.ID.PRT:lens
F	17.6	PRT:[diaphram] [camera]

+Paragraph 3

Behind the cornea is the lens. [structure]

S	18	OBJ:lens LOC:behind REL.OBJ: cornea

The lens consists of a number of transparent layers arranged much like the layers of an onion. [structure]

S	19	PRC:consist PAT:lens DET:definite NUM:singular REL.PROP:19.1
S	19.1	PRT:[lens] [layer]
S	19.2	OBJ:layer ATT:transparent NUM:plural
S	19.3	ACT:arrange OBJ:19.2, PROX.ATT:*much like* [19.4]
S	19.4	ACT:arrange OBJ:onion

STATE.ID.PRT:layer

These layers are held in place in a sac. [structure]

S	20	ACT:hold
		OBJ:layer
		DET:definite
		NUM:plural
		LOC:in a sac
		RSLT: 20.1
		VCE:passive

/picture 13

After light passes through the pupil, it next encounters the lens. [energy]

E	21.1	ACT:move
		OBJ:layer
		TRTH:Negative
E	21.2	ACT:pass
		AGT:light
		RSLT:21.2
		DIR:*through*
E	21.2	OBJ:light
		LOC:pupil
		TRTH:neg
E	21.3	ACT:encounter
		AGT:it
		REL.OBJ:lens
		RSLT:21.4
E	21.4	OBJ:light
		LOC:lens
E	21.5	ORD.LOC:[21.0]
		[21.2]
		after
		next

Although most of the refraction or bending of light is done by the cornea, some remaining focusing is accomplished by varying the thickness and hence the refractive power of the lens. [function]

F	22	EQUIV.TEM:*as*
		[22.1]
		[22.2]
F	22.1	PRC:refraction
		PAT:light
		DEG:most
F	22.2	ACT:PASS
		AGT:it(light)
		SRC.LOC:22.3
		RSLT.LOC:22.4

F	22.3	OBJ:light LOC:air DIR:from
F	22.4	OBJ:light LOC:cornea DIR:into
F	22.5	ACT:bend OBJ:light ASP:continuing
F	22.6	EQUIV:[refraction] [22.5]
F	22.7	PRC:focus PRC.ATT:remaining DEG:some PRT.PRC:22.8
F	22.8	ACT:vary OBJ:lens ATT:thickness ASP:continuing
F	22.9	ACT:elided (vary) OBJ:lens ATT:refractive power
F	22.10.	COND: [22.8] [22.9]

This process is called accommodation. [other]

	23	EQUIV: "called" [this process] [accomodation]

In this process, the ciliary muscles tighten and thereby increase the focusing power of the eye. [function]

F	24	PRC:process DET:definite NUM:singular PRT.PRC:24.1
F	24.1	PRC:focus PAT:eye PRC.ATT:power ASP:continuing DEG:increase
F	24.2	ACT:tighten AGT:24.3 RSLT:24.1
F	24.3	OBJ:muscle

		STATE.ID.ATT:ciliary
		DET:definite
		NUM:plural
F	24.4	COND:
		[24.2]
		[24.1]

The curvature of the lens is increased by this tightening of muscles. [function]

F	25	ACT:tighten
		OBJ:muscle
		DET:generic
		NUM:plural
F	25.1	OBJ:lens
		DET:definite
		NUM:singular
		ATT:curvature
		DGR:increase
F	25.2	COND: *this*
		[24.3]
		[25.1]

Accommodation enables us to keep an image on the retina sharp as we look at objects located at different distances. [function]

F	26	COND:
		[accomodation]
		[26.3]
F	26.1	ACT:keep
		AGT:us
		OBJ:image
		DET:generic
		NUM:singular
		ATT: sharp
F	26.2	ACT:look
		AGT:we
		OBJ:objects
		ATT:distance
		MEAS:different
		NUM:plural
F	26.3	EQUIV.TEM:*as*
		[26.1]
		[26.2]

+Paragraph 4

The vitreous humor, a clear jellylike substance, is located in the middle of the eyeball. [structure]

S	27	OBJ:humor

		STATE.ID.ATT:vitreo us LOC:27.2 DIR:in
S	27.1	OBJ:substance ATT:clear, jellylike
S	27.2	PRT: [eyeball] [middle]
S	27.3	IDENT: [27.0] [27.1]

The purpose of this substance is to maintain the shape of the eye, and also to provide a medium with a similar refractive index to that of the lens, so that no further refraction of light occurs. [function]

F	28	PRC:maintain PAT:27.2 R.OBJ:eye ATT:shape DET:definite NUM:singular
F	28.1	OBJ:substance DET:defininte NUM:singular
F	28.2	PRC:provide R.OBJ:28.3
F	28.3	OBJ:medium ATT:refractive index PROX.ATT: 28.4
F	28.4	OBJ:lens ATT:that (refract index)
F	28.5	COND:"so that" [25.6] [28.7]
F	28.6	PRC:refraction OBJ:light DGR:further TRTH: neg
F	28.7	AND [28.0] [28.2]

After the light exits from the lens it passes through the vitreous humor. [energy]

F	29	ORD.LOC:"after" [29.1]

		[29.2]
F	29.1	ACT:exit AGT:light SRC.LOC:lens DET:definite NUM:singular
F	29.2	ACT:pass AGT:it(light) R.OBJ:vitreous humor DIR:through
F	29.3	OBJ:vitreous humor DET:definite NUM:singular DET:definite NUM: singular

The light is then focused on the inside surface of the ball. [energy]

E	30	ACT:focus OBJ:(light) R.OBJ:30.1 LOC:on DIR:inside
E	30.1	OBJ:ball STATE.ID.PRT:surfac e DET:definite NUM:singular
E	30.2	ORD:TEM [29.2] [30.0]

The retina is an area of the inside surface of the eyeball that surrounds nearly 200 degrees as measured on the circumference of a circle. [structure]

S	31	PRC:surround PAT:31.1 R.OBJ:eyeball LOC:inside DEG:nearly MEAS:200 UNIT:degree
S	31.1	ISA: [retina] [area]
S	31.2	PRT: [eyeball]

		[retina]
		LOC:inside
		surface
S	**31.3**	ACT:measure
		INST:31.4
S	**31.4**	OBJ:circle
		ATT:circumfrance
S	**31.5**	COND:*as*
		[3.10]
		[31.3]

The photoreceptors and their neural support are embedded in the retina. [structure]

S	**32**	PRC:embed
		PAT:photoreceptor
		R.OBJ:retina
		LOC:in
S	**32.1**	PRC:embed
		PAT:32.2
		R.OBJ:retina
		LOC:in
S	**32.2**	OBJ:support
		ATT:neural
S	**32.3**	POSS:*their*
		[photoreceptor]
		[32.2]

The fovea is an area of the retina located directly behind the lens. [structure]

S	**33**	OBJ:fovea
		ATT:area
		LOC:lens
		DIR:behind
S	**33.1**	PART:
		[retina]
		[fovea]

/picture 14

It is the most sensitive portion of the retina to detecting light patterns. [function]

F	**34**	OBJ:fovea
		ATT:sensitive
		DEG:most
		R.PRC:34.2
F	**34.1**	PRT:[fovea]
		[fovea]
		[retina]
F	**34.2**	PRC:pattern

The blind spot, or optic disk, is an area on the retina that covers several degrees.
[structure]

S	35	EQUIV:
o		[blind spot]
		[optic disk]
S	35.1	PRT:
		[blind spot]
		[retina]
S	35.1	OBJ:blind
		spot
		ATT:area
		MEAS:several
		UNIT:degree
		LOC:on
		R.OBJ:retina

It has no receptors because the nerve connections from the receptors exit from the eye at this point to form the optic nerve tract connecting to higher centers in the brain.
[structure]

S	36	POSS:
		[it]
		[receptors]
		TRTH:negative
S	36.1	CAUS:
		[36.2]
		[36.0]
S	36.2	ACT:exit
		AGT:36.3
		R.OBJ:eye
		LOC:at
		this
		point
		GOAL:36.4
S	36.3	OBJ:connection
		ATT:nerve
		SRC.LOC:receptors
		DIR:from
S	36.4	ACT:form
		AGT:36.3
		RSLT:36.5
S	36.5	OBJ:tract
		ATT:optic nerve
S	36.6	ACT:connect
		OBJ:36.5
		R.OBJ:36.7
S	36.7	OBJ:center
		NUM:plural
		ATT:higher

		LOC:in
		R.OBJ: the brain

+Paragraph 5

The photoreceptors convert light energy into neural energy. [function]

F	**37**	ACT:convert
		AGT:photoreceptor
		NUM:plural
		SRC:37.1
		RSLT:37.2
F	**37.1**	OBJ:energy
		ATT:light
F	**37.2**	OBJ:energy
		ATT:neural

They are embedded in the back of the retina, with all of the neural connections and blood supply in front of them. [structure]

S	**38**	PRC:embed
		PAT:they(photorece
		ptors)
		R.OBJ:retina
		LOC:in back of
S	**38.1**	OBJ:connection
		NUM:all
		ATT:neural
		LOC:in front of
		R.OBJ:them
S	**38.2**	OBJ:supply
		NUM:all
		ATT:blood
		LOC:in front of
		R.OBJ:them
		(photoreceptors)
S	**38.3**	AND
		[38.1]
		[38.2]

/picture 21

Consequently, the light passes through all of the supporting structures before reaching the receptors. [energy]

E	**39**	CAU:
		[ ]
		[39.1]
E	**39.1**	ACT:pass
		AGT:light
		R.OBJ:39.3

		DIR:through RSLT:39.5
E	39.2	OBJ:light LOC:supporting structures
E	39.3	OBJ:structure ATT:supporting NUM:all
E	39.4	ACT:reach AGT:light LOC:receptors
E	39.5	ORD:TEM: [39.2] [39.4]

There are two kinds of photoreceptors embedded in the back of the retina called rods and cones. [structure]

S	40	CAT: [photoreceptors] [40.3, 40.4]
S	40.1	PRT: [photoreceptors] [40.2]
S	40.2	OBJ:system ATT:human, visual
S	40.3	OBJ:rod NUM:plural
S	40.4	OBJ:cone NUM:plural

Rods are the rod-shaped receptors that are primarily responsible for vision in low levels of illumination. [function]

F	41	IDENT: [rods] [41.1]
F	41.1	OBJ:receptor ATT:rod-shaped NUM:plural
F	41.2	PRC:vision PAT:41.1
F	41.3	PRC:illumination ATT:level DEG:low
F	41.4	COND:"responsible for" [41.3] [41.2]

The rod system is extremely sensitive in the dark, but cannot resolve fine details. [function]

F 42 OBJ:system
ATT:rod
F 42.1 OBJ:42.0
ATT:sensitive
DEG:extremely
F 42.2 COND:
[darkness]
[42.1]
F 42.3 ACT:resolve
AGT:42.1
OBJ:detail
MOD:can
NUM:plural
ATT:fine
TRTH:neg

Cones are cone-shaped receptors that are primarily responsible for vision in high levels of illumination. [function]

F 43 IDENT:
[cones]
[43.1]
F 43.1 OBJ:receptor
ATT:cone-shaped
NUM:plural
F 43.2 PRC:vision
PAT:43.1
F 43.3 PRC:illumination
ATT:level
DEG:high
F 43.4 COND:"responsible for"
[43.3]
[43.2]

The cone system is responsible for detail vision and color vision. [function]

F 44 CAU:"responsible"
[44.1]
[44.4]
F 44.1 OBJ:system
ATT:cone
F 44.2 PRC:vision
ATT:detail
F 44.3 PRC:vision
ATT:color
F 44.4 AND:

[44.2]
[44.3]

+Paragraph 6

There are about 120 million rods and nearly seven million cones. [structure]

S	**45**	OBJ:rod
		NUM:120
		UNIT:million
		DGR:about
S	**45.1**	OBJ:cone
		NUM:7
		UNIT:million
		DGR:nearly
S	**45.2**	AND
		[45.0]
		[45.1]

Most of the cones are contained in the center of the fovea, which covers an area of perhaps 1 degree. [structure]

S	**46**	PRC:contain
		PAT:46.3
		R.OBJ:46.1
		VC:passive
S	**46.1**	OBJ:cone
		DET:generic
		NUM:plural
		DEG:most
S	**46.2**	PRC:cover
		PAT:46.3
		ATT:area
		DET:definite
		NUM:one
		DGR:perhaps
		VAL:1
		MEAS:degree
S	**46.3**	PRT:
		[center]
		[fovea]

The number and proportion of cones falls off rapidly with many fewer cones present beyond 10 degrees. [structure]

S	**47**	ACT:fall off
		OBJ:cone
		ATT:number
		DGR:rapidly
		RSLT:47.3
S	**47.1**	ACT:fall

off
OBJ:cone
ATT:proportion
DGR:rapidly
RSLT:47.3

S 47.2 AND
[47.0]
[47.1]

S 47.3 OBJ:cone
DGR:many
NUM:fewer
LOC:beyond
MEAS:10
MEAS: degee

On the other hand, there are no rods in the center of the fovea. [structure]

S 48 ACT:find
OBJ:rod
NUM:none
R.OBJ:48.1

S 48.1 OBJ:fovea
LOC:in the center

They reach their highest frequency at about 16 degrees on either side, with decreasing numbers out to about 100 degrees at the edge of the retina on either side of the fovea. [structure]

S 49 ACT:reach
OBJ:they(rods)
ATT:frequency
DEG:highest
DIR:on either side
MEAS:16
UNIT:degree

S 49.1 OBJ:empty (rods)
"with"
ATT:number
DEG:decreasing
DIR:out
MEAS:100
UNIT:degree
LOC:49.2

S 49.2 OBJ:retina
ATT:edge
LOC:fovea
DIR:on either side

S 49.3 AND: "with"
[49.0]
[49.1]

+Paragraph 7

The photoreceptors synapse onto bipolar and then ganglion cells. [structure]

S	50	ACT:synapse OBJ:photoreceptor NUM:plural REL.OBJ:50.1, 50.2 DIR:onto
S	50.1	OBJ:cell CAT(STATE.IDENT.A TT):bipolar NUM:plural
S	50.2	OBJ:cell CAT(STATE.IDENT.A TT):ganglion NUM:plural
S	50.3	ORD: [50.1] [50.2]

/picture 22

Electrical signals start in the photoreceptors and travel to the brain via these synaptic connections. [energy]

E	51	ACT:signal ACT.ID.ATT:electrica l ASP:iterative, inceptive REL.ACT:51.1 NUM:plural LOC:in the photoreceptors
E	51.1	ACT:travel INST:51.2 GOAL.LOC:brain DIR:to
E	51.2	OBJ:connection STATE.IDENT.ATT:sy naptic DET:definite NUM:plural
E	51.3	ORD.TEM: *start* [50.1] [51.1]

In the fovea, usually only one cone is connected to one bipolar cell. [structure]

S	52	ACT:connect OBJ:cone NUM:one MOD:usually REL.OBJ:52.1 LOC: 52.2
S	52.1	OBJ:cell STATE.IDENT.ATT:bi polar NUM:one
S	52.2	OBJ:fovea DET:definite NUM:singular LOC:in

Outside of the fovea there will be many photoreceptors connected to one bipolar cell. [structure]

S	53	ACT:connect OBJ:receptor NUM:many REL.OBJ:53.1 LOC:53.2
S	53.1	OBJ:bipolar NUM:one
S	53.2	OBJ:fovea DET:definite NUM:singular LOC:outside

The number of photoreceptors converging onto a single bipolar cell approaches a convergence of hundreds of rods onto one bipolar beyond 20 degrees into the periphery. [structure]

S	54	PROX.IDENT *approaches* [54.1] [54.3]
S	54.1	PRC:converge PAT:photoreceptor PAT.ID.ATT:number R.OBJ:54.2 DIR: onto NUM:plural ASP:continuing
S	54.2	OBJ:cell STATE.ID.ATT:bipola r NUM:single DET:generic

S	54.3	PRC:converge PAT:rod R.OBJ:54.4 DIR:onto NUM:plural UNIT: hundred
S	54.4	OBJ:cell STATE.ID.ATT:bipola r NUM:one
S	54.5	OBJ:elided (fovea) LOC:periphery DEG:beyond DIR:into MEAS:20 UNIT: degree
S	54.6	COND: [54.3] [54.5]

This great pooling of receptors onto a single bipolar cell in the periphery means that any particular bipolar cell cannot determine which of its many receptors has been stimulated by light. [structure & function]

S	55	ACT:pool OBJ:receptor NUM:plural ACT.ATT:great DIR:onto R.OBJ:55.1 RSLT:55.2 ASPT:iterative
S	55.1	OBJ:cell STATE.ID.ATT:bipola r NUM:singular MEAS:single DET:generic LOC:periphery DIR:in
F	55.2	ACT:mean THM:55.3, 55.4
F	55.3	ACT:determine AGT:bipolar DET:any particular MOD:can TRTH: neg R.ACT:55.4

F	**55.4**	ACT:stimulate AGT:light OBJ:receptor OBJ.CAT:which NUM: plural MEAS:many TNS:past
F	**55.5**	POSS [bipolar] [receptor]
F	**55.6**	IDENT:"this" [55.1] [55.0]

On the other hand, in the center of the fovea there is very little pooling, resulting in virtually perfect specificity of excitation. [function]

F	**56**	ACT:pool DGR:little ASPT:iterative LOC:56.1 RST:56.2
F	**56.1**	OBJ:fovea DET:definite NUM:singular LOC:center DIR:in
F	**56.2**	PRC:excitation ATT:specificity ATT:perfect DGR:virtually
F	**56.3**	COND:"resulting in" [56.0] [56.2]

+Paragraph 8

Each bipolar cell is connected to a ganglion cell via a second synapse. [structure]

S	**57**	ACT:connect INST:synapse INST.ATT:second OBJ:57.1 R.OBJ:57.2
S	**57.1**	OBJ:cell STATE.IDENT.ATT:bi polar DET:each
S	**57.2**	OBJ:cell

STATE.IDENT.ATT:ga
nglion
DET:generic
NUM:singular

The ganglion cell has an elongated body that forms one of the fibers of the optic nerve. [structure]

S	58	OBJ:cell STATE.IDENT.ATT:ga nglion DET:defninite NUM:singular
S	58.1	PRT: [cell] [body]
S	58.2	OBJ:body ATT:enlongated
S	58.3	ACT:form AGT:that(body) RSLT:58.4, 58.5
S	58.4	OBJ:fiber NUM:plural MEAS:one
S	58.5	PRT: [fiber] [optic nerve]

There are less than one million ganglion cells leaving the retina in this manner. [structure]

S	59	ACT:leave AGT:59.1 R.OBJ:retina NUM:less than MEAS:one UNIT:million ACT.ATT:59.2
S	59.1	OBJ:cell ATT:ganglion
S	59.2	PROX.ATT:*in this manner* [59.3] [59.0]

Again, in the fovea each bipolar cell generally connects to one ganglion, whereas in the periphery a number of bipolars will converge on a single ganglion cell. [structure]

S	60	ACT:connect AGT:60.1 OBJ:ganlion

		NUM:one DGR:generally LOC: 60.2 ASPT:continuing
S	60.1	OBJ:cell STATE.ID.ATT:bipola r NUM:singular DET:each
S	60.2	OBJ:fovea DET:definite NUM:singular LOC:in
S	60.3	ACT:converge AGT:60.4 OBJ:60.5 DIR:on
S	60.4	OBJ:bipolar NUM:plural *a number* LOC:periphery DIR:in TNS: future
S	60.5	OBJ:cell STATE.ID.ATT:gangli on NUM:single

The two remaining types of retinal cells are horizontal and amacrine cells. [structure]

S	61	ACT:call OBJ:cell NUM:plural ATT:remaining NUM:2
S	61.1	OBJ:cell STATE.ID.ATT:horizo ntal
S	61.2	OBJ:cell STATE.ID.ATT:amacr ine
S	61.3	AND: [61.2] [61.3]

Horizontal cells connect receptors to other receptors. [structure]

S	62	ACT:connect AGT:62.1 OBJ:receptor

		NUM:plural
		DIR:to
		R.OBJ: 62.2
S	62.1	OBJ:cell
		NUM:plural
		ATT:horizontal
S	62.2	OBJ:receptor
		NUM:plural
		DET:other

Amacrine cells connect ganglion cells to other ganglion cells and also bipolar cells to other bipolar cells. [structure]

S	63	ACT:connect
		AGT:63.1
		OBJ:63.2
		DIR:to
		R.OBJ:63.3
S	63.1	OBJ:cell
		NUM:plural
		ATT:amacrine
S	63.2	OBJ:cell
		NUM:plural
		ATT:ganglion
S	63.3	OBJ:cell
		NUM:plural
		ATT:ganglion
		DET:other
S	63.4	ACT:connect
		AGT:63.1
		OBJ:63.5
		DIR:to
		R.OBJ:63.6
S	63.5	OBJ:cell
		NUM:plural
		ATT:bipolar
S	63.6	OBJ:cell
		NUM:plural
		ATT:bipolar
		DET:other
S	63.7	AND:
		[63.0]
		[63.4]

/picture 23

These cells do not transmit signals towards the brain, but instead they transmit and pool signals laterally across the retina. [function]

F	64	ACT:transmit

		AGT:cell
		DET: these
		NUM: plural
		THM:64.1
		DIR:towards
		LOC:the brain
		TRTH:neg
F	64.1	ACT:signal
		ASP:iterative
F	64.2	ACT:transmit
		AGT:they
		THM:64.4
		DIR:laterally
		RSLT:64.5
F	64.3	ACT:pool
		AGT:they
		THM:64.4
		DIR:laterally
		RSLT:64.5
F	64.4	ACT:signal
		ASP:iterative
F	64.5	ACT:signal
		DIR:across
		LOC:the
		retina
F	64.6	AND:
		[64.2]
		[64.3]

Thus, neural signals flow both directly towards the brain and laterally across the retina before going to the brain. [energy]

E	65	ACT:flow
		THM:65.1
		DIR:directly towards
		R.OBJ:the brain
E	65.1	ACT:signal
		ACT.ID.ATT:neural
		ASP:iterative
E	65.2	ACT:flow
		THM:65.3
		DIR:laterally
		RSLT:65.4
E	65.3	ACT:signal
		ASP:iterative
E	65.4	ACT:signal
		DIR:across
		LOC:the retina
E	65.5	ACT:go

		OBJ:elided (signals)
		DIR:to
		LOC:the brain
E	65.6	AND:
		[65.0]
		[65.2]
E	65.7	ORD.TEM:"before"
		[65.2]
		[65.5]

+Paragraph 9

We still do not completely understand the details concerning how the energy in photons is transduced into electrical activity in neurons. [other]

	66	PRC:understand
		PAT:we
		THM:66.1, 66.2,
		66.3
		ASPT:continuing
		DGR:completely
		TRTH:neg
	66.1	ACT:concern
		PRT:details
		THM:66.2
	66.2	ACT:how
		THM:66.3
	66.3	ACT:transduce
		OBJ:60.4
		DIR:into
		RSLT:66.5
	66.4	PRT:
		[photons]
		[energy]
	66.5	ACT:activity
		ATT:electrical
		LOC:in neurons

The first stage requires absorption of a photon by a molecule of photopigment in a photoreceptor. [function]

F	67	ACT:absorbe
		AGT:67.1
		OBJ:photon
		DET:generic
		NUM:singular
F	67.1	OBJ:molecule
		ATT:photopigment
		LOC:photoreceptor
F	67.2	PRC:stage

		MOD:necessity
		requires
		PRT:67.0
F	67.3	ORD:TEM:
		[67.2]
		[empty]

The photopigment in rod receptors is called rhodopsin, or visual purple. [structure]

S	68	IDENT:*called*
		[68.1]
		[rhodopsin]
S	68.1	OBJ:photopigment
		DET:definite
		NUM:singular
		LOC: in
		R.OBJ: 68.2
S	68.2	OBJ:receptor
		DET:definite
		NUM:plural
		STATE.ID.ATT:rod
S	68.3	EQUIV:
		[rhodopsin]
		[visual purple]

/picture 24

Each of the four million or so molecules of rhodopsin in each rod will undergo a molecular change upon the absorption of a photon. [function]

F	69	COND:
		[69.5]
		[69.1]
F	69.1	ACT:change
		STATE.ID.ATT:molec
		ular
		OBJ:69.2
		ASP:incept
F	69.2	OBJ:molecule
		DET:generic *each*
		NUM:plural
		MEAS:4
		UNIT: million
		DEG: *or so*
F	69.3	PRT:
		[molecules]
		[rhodopsin]
F	69.4	OBJ:rod
		DET:generic *each*

F	69.5	PRC:absorb
		PAT:it (molecule)
		R.OBJ:photon
		DET:generic
		NUM: single

This change occurs almost instantaneously with absorption.

F	70	ACT:change
		DET:*this* [ ]
		[69.1]
		ASP:incept
F	70.1	ACT:absorption
F	70.2	EQUIV.TEM:
		[70.0]
		[70.1]
		DEG:almost

The products of this change are capable of producing a neural charge. [function]

F	71	CAUS
		[71.2]
		[71.1]
F	71.1	ACT:produce
		AGT:71.2
		MOD:capable
		RSLT:71.3
F	71.2	ACT:change
		DET:*this* [69.1]
		RSLT:products
F	71.3	OBJ:charge
		STATE.ID.ATT:neural

This charge flows across the synapse between the photoreceptor and the bipolar cell it connects to. [energy]

E	72	ACT:flow
		OBJ:charge
		DET:definite *this*
		DIR:across
		LOC:72.1
E	72.1	OBJ:synapse
		LOC:between
		R.OBJ:72.2, 72.3
E	72.2	OBJ:photoreceptor
		DET: definite
E	72.3	OBJ:cell
		ATT:bipolar
E	72.4	ACT:connect
		OBJ:72.2
		R.OBJ:72.3

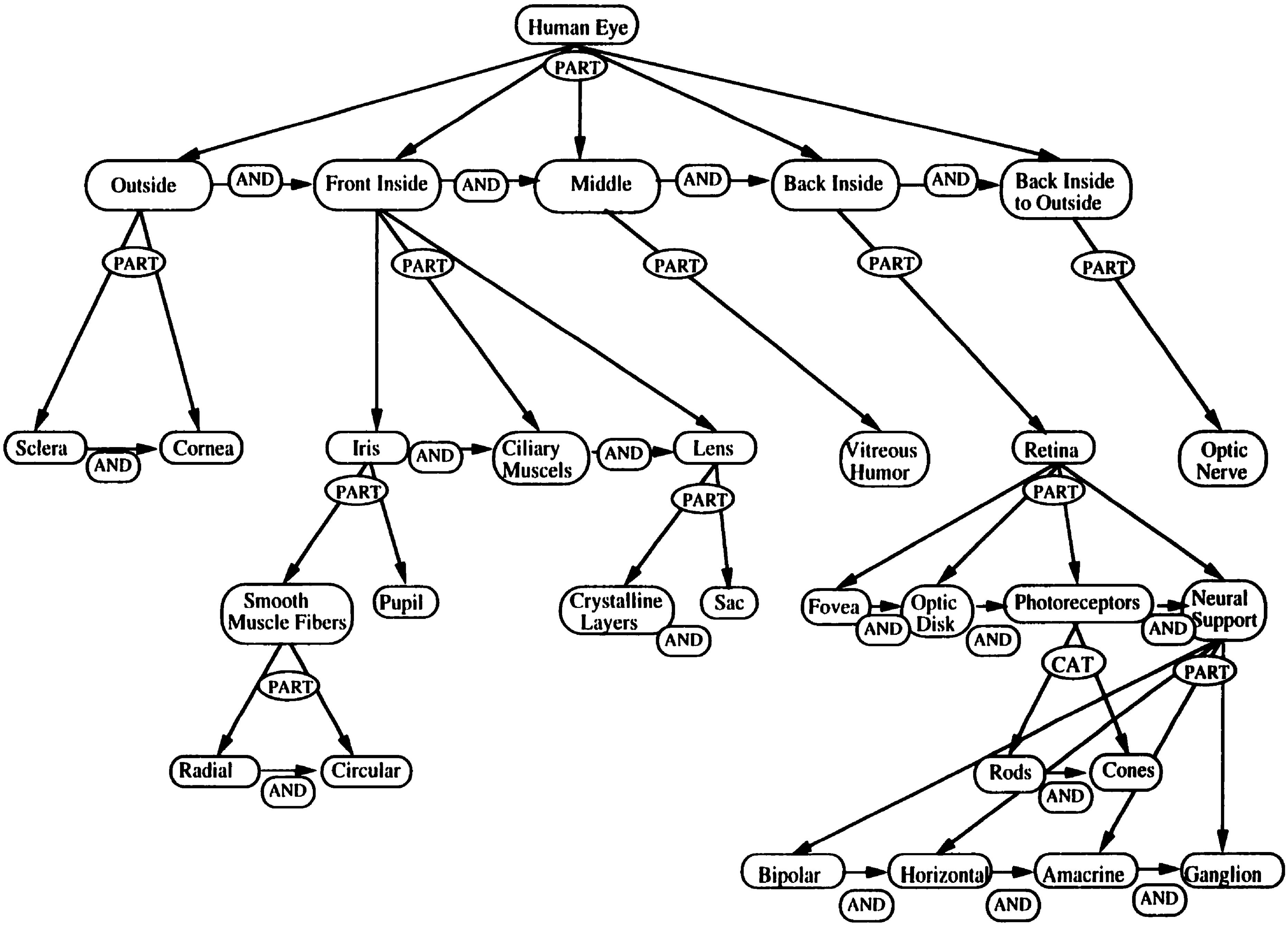
Note. / = locations of forced exposure to an illustration for participants in the Controlled Access condition; + = locations where a new paragraph begins; [] = information type for each sentence; E = energy; F = function; S = structure.

ACT = action; AGT = agent; ASP = aspect; ATT = attribute; CAUS = cause; COND = condition; DEG = degree; DET = determiner; DIR = direction; EQUIV = equivalence; IDENT = identity; LOC = location; MEAS = measure; MOD = modality; NUM = number; OBJ = object, PAT = patient; PRC = process; PRT = part; R.OBJ = related object; RSLT= result; TEM = temporal; THM = theme; TRTH = truth value; UNIT = unit

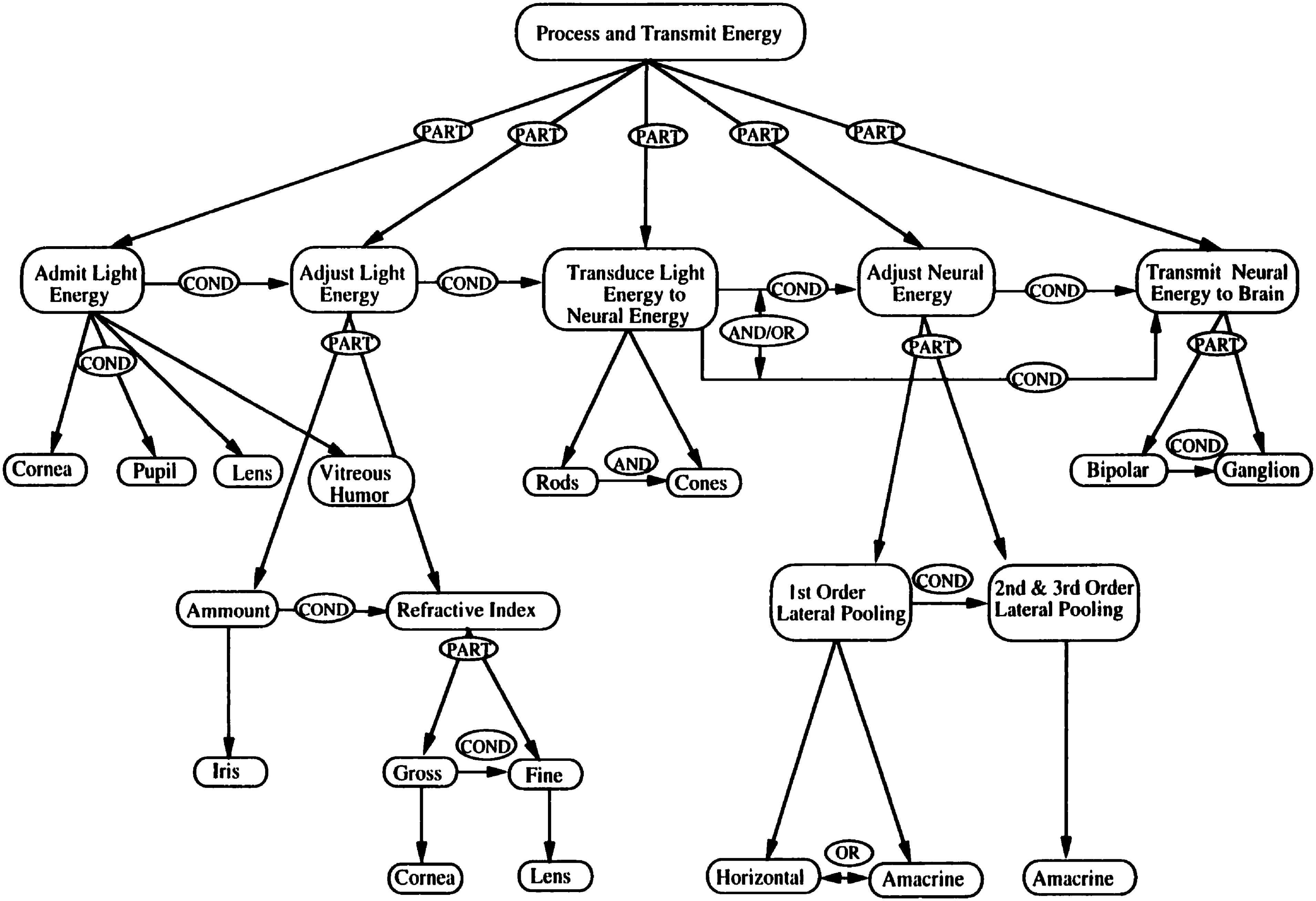
Appendix G

Mental Models of the Visual System

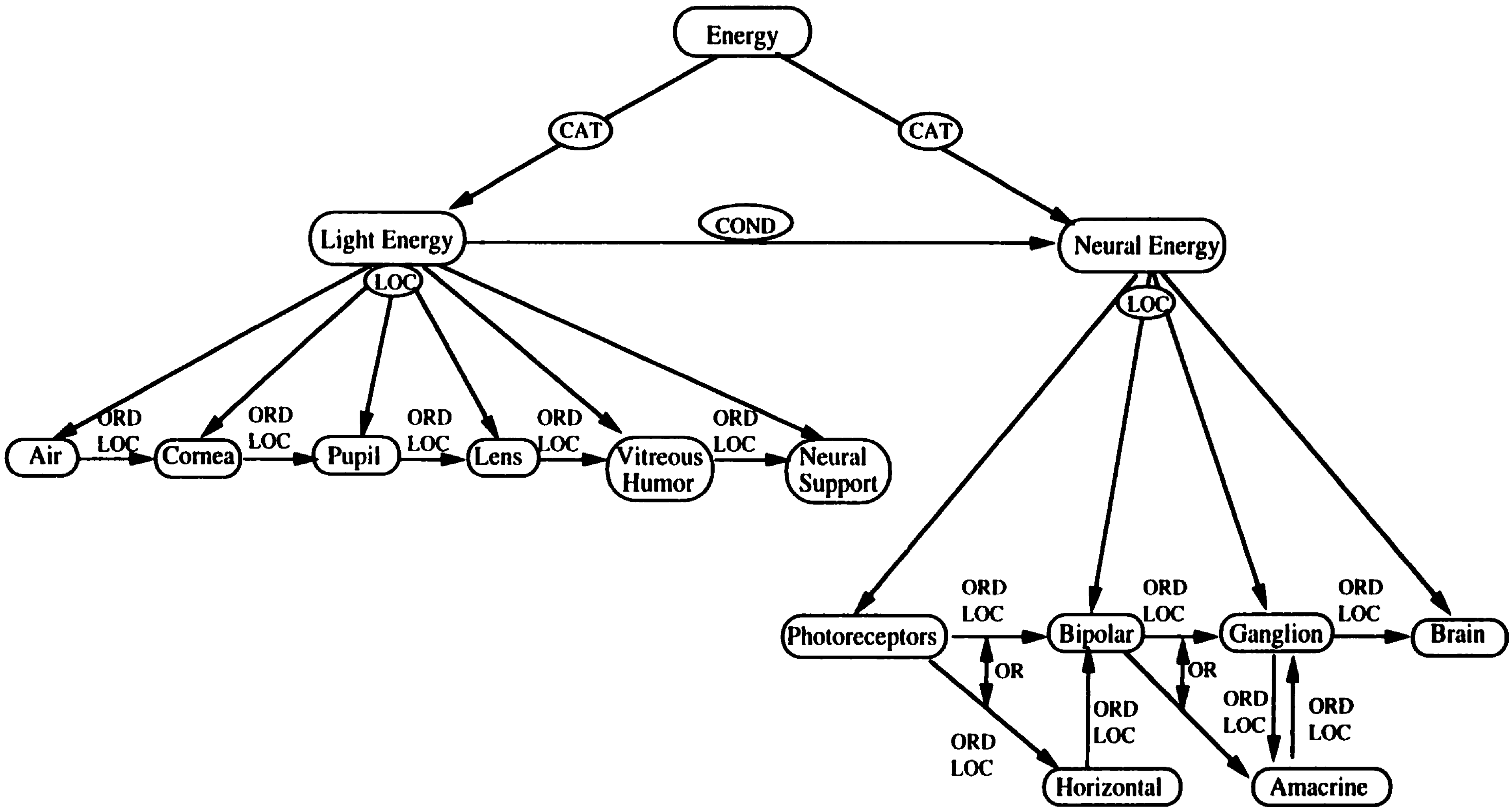
Structural Organization of Components



Functional Organization of Components



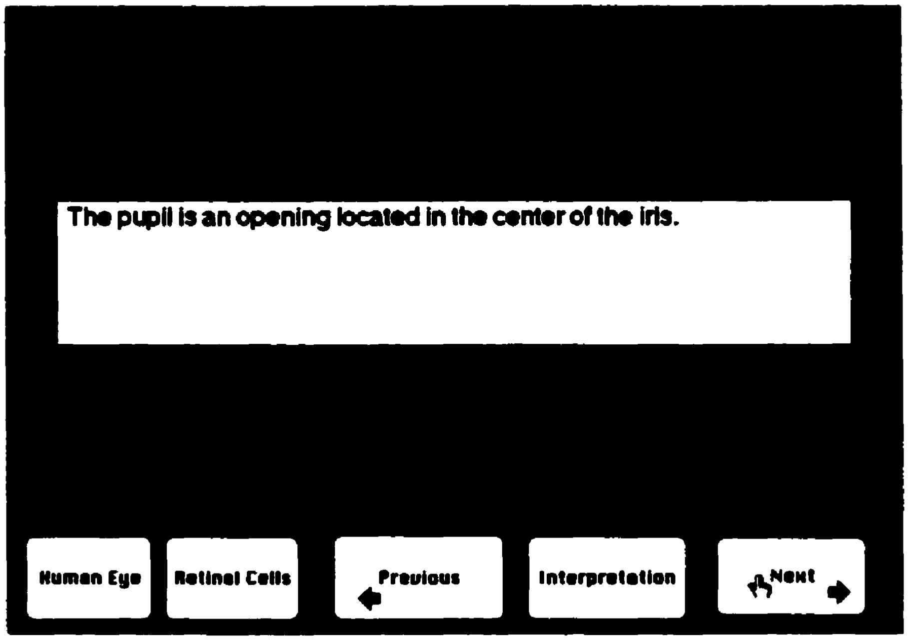
Organization of Components by Energy Path



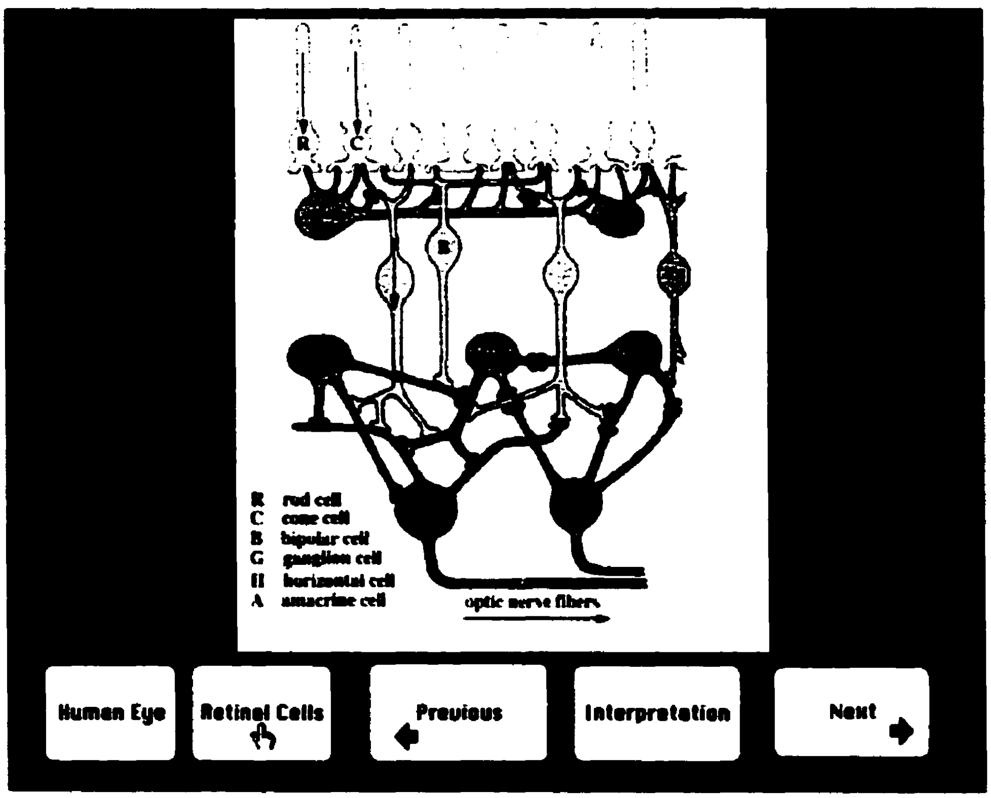
Appendix H

Screen Shots of Presentation Environment

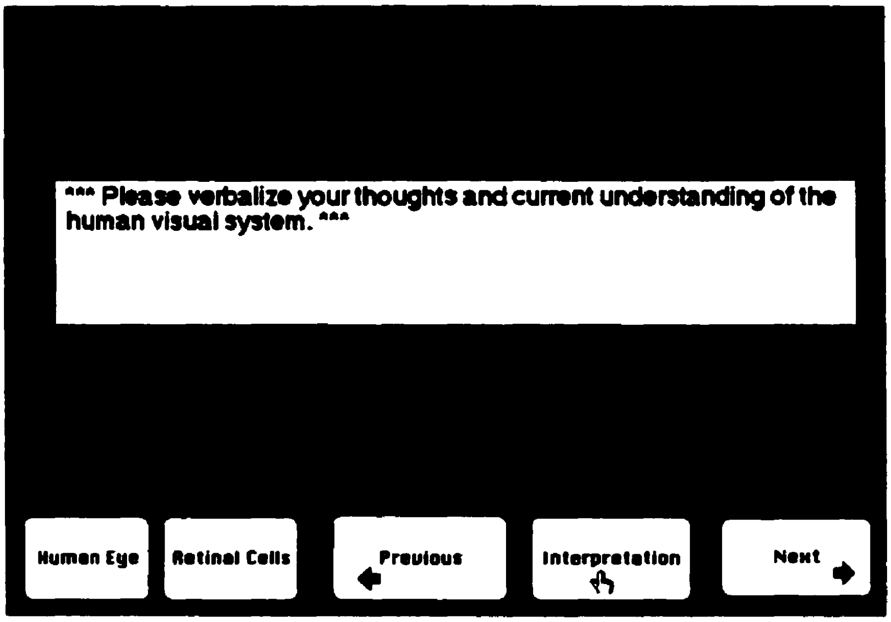
(a) Example of sentence presentation screen.



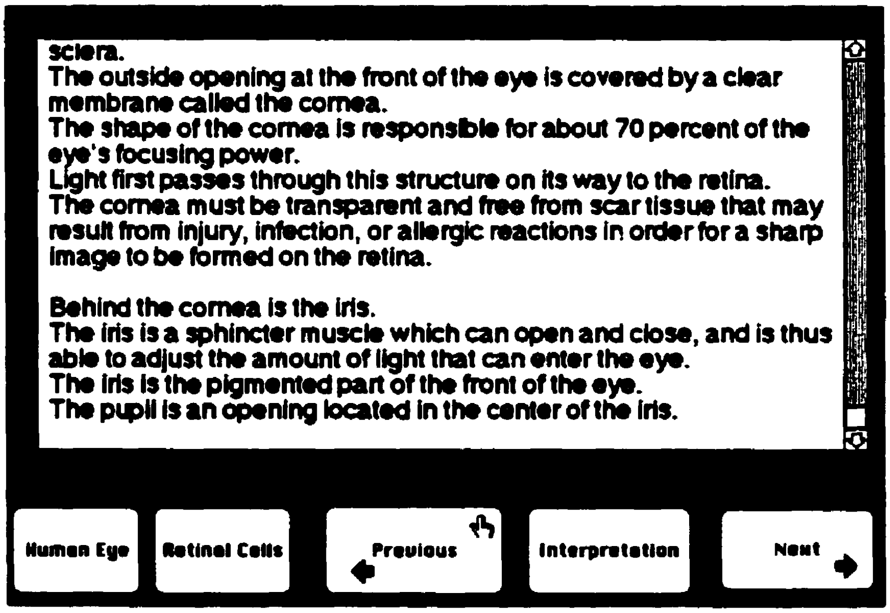
(b) Example of illustration presentation screen.



(c) Example of on-line interpretation screen.



(d) Example of text look-back screen.



Appendix I

Post-Input Comprehension Questions

Questions That Require Integration of Information Described in the Text Only

Question 1:
Based on your understanding of the visual system explain why humans see better in light than in dark conditions?

Scoring Question 1
(1point) MODEL 1 - explanation refers to a "special amount of light"
(1point)MODEL 2 - more light = more information
(2point)MODEL 3 - receptor specialization to different wavelengths
(4point)MODEL 4 - structural location of receptors
(5point)MODEL 5 - receptor specialization and location

Example of a Response From a Participant in Illustrations Only Condition
[Score for Question 1 = Model 2 = 1 point]
Ah this is I guess just a guess
more light rays can enter ah the eye
the more light there can be umm absorbed by the eye
the more information can be given passed on to the nerves.
and so we can interpret more.
It's kind of like a puzzle.
If ah the more pieces that we can turn around
and so the more light that there is
the the the clearer the picture becomes.

Question 3:
Explain the process of accomodation.

Scoring Question 3
General Answer:
ciliary muscles expand and contract to bend lens to adust to different distances (score = 2)

Specific Answer:
ciliary muscles tighten->lens is less convex (1) -> focal point shifted foreward for objects at far

distance (1) ; ciliary muscles relax-> lens more convex (1) ->focal point shifted backward for objects at close distance (1) (total score = 4)

Example of a Response From a Participant in the Text Only Condition

[Question 2 Score = 2]

Ya because of accomodation

which wasss done by the lens. (.5)

For different distances?

Ya

the cornea tries to get it towards the foeva

I suppose

and then the lens really like gets it right on.

And then by stretching it out

or skwishing it **(1)**

the muscles around the lens can put tension or pressure on

and we focus that way. (.5)

Question 7:

Explain how the human eye can adjust to different lighting conditions.

Scoring Question 7

circular iris muscles contract -> pupil contracts -> less light admitted (3)

radial iris muscles contract -> pupil dilates -> more light admitted (3) (total score = 6)

Example of a Response From a Participant in the Text Only Condition

[Question 7 Score = 3]

Mk.

I would think that the pupil would dilate **(1)**

so the radial fibers in the iris would umm contract. **(1)**

And that that would allow more light in. **(1)**

And I would also think that... OK what was the question again?

How

E: How it compensates? Ya or how it adjusts to different lighting conditions.

How it adjusts (sigh)

Because I was thinking as I was reading that that this isn't really how the eye would adjust

but this is how you could adjust to it.
By looking at things through through the corner of your eye sort of indirectly
because there's well first of all there's the blind spot
but secondly most of the rod zones are around the periphery.
So I would think that if look at something directly
you wouldn't see as well.
Which I've noticed when I look at stars I don't see them very well ().
Umm what else?
That's all I can think of really.
Is the dilating pupil.
That's about it.

Questions That Require Integration of Information Described in both the Text and Illustrations

Question 2:
Explain how light enters the eye and is adjusted and directed to the retina.

Scoring for Question 2 (possible total score = 6)

Energy Path	Function
Object -> Cornea (.5)	Cornea: admits and refracts (gross) light (.5)
	Iris regulates size of Pupil (.5)
Cornea -> Pupil (.5)	Pulil admits regulated amount of light (.5)
	Ciliary Muscles adjust lens (.5)
Pupil -> Lens	Lens admits and regracts (fine) light (.5)
Lens -> Vitreous Humor (.5)	Vitreous Humor admits and maintains refraction of light (.5)
Vitreous Humor -> Retina (.5)	Retina recieves light where most light is focused on the fovea (.5)

Example of a Response From a Participant in the Free Access Condition

[Question 2 Score = 4]
Ok umm
There's ah the light passes through the cornea first. (.5)

And the cornea does 70% of the bending of the light. (.5)
() as a refractive index
that means as as when light hits it
it bends
and it's focused
so the cornea does that part ah 70% of the focusing
and the light then goes through the lens (.5)
which focuses it more (.5)
and umm directs it to to the back of the retina at the fovea. (1)
And it goes through the the liquid (.5)
and it isn't it isn't ah distorted or refracted more. (.5)

Question 5:
Explain in as much detail as you can how light is transduced into neural energy in receptor cells and how the neural energy gets pooled and transmitted to the brain.

Scoring for Question 5
light is absorbed by photopigment (1) in photoreceptors (1) -> transduction from light energy into electrochemical charge (1), travels from photoreceptor to bipolar (1), from bipolar to ganglion(1), from ganglion to brain(1), optional lateral pooling across receptors by horizontal cell (1), lateral pooling across bipolar /ganglion cells by amacrine cells (1) (total score = 8)

Example of a Response From a Participant in the Illustrations Only Condition
[Question 5 Score = 4.5]
Umm elelctrical energy?
Ah ok.
Well ok my very very superficial understanding would be umm inside the receptors
that you have a ah a specific chemical I guess (1)
so when the light hits the chemical (1)
you probably would have a electron excitement
or whatever ah this change would
you know I think of it more as electrochemical rather than electrical right (1)
Through the you know I guess this fluctuation would travel through the nerves (.5)

which I guess are specifically designed to to carry this ele the electrical impulse chemically through through throuh the length of their bodies.
And then you know it's well is it phosphorous?
I can't remember
well anyhow it starts with a p umm one of the elements
and well umm
and it ah
so it travels one one atom at a time
and then at the end of the cell it umm
I guess it umm a different chemical from the one I remember
it passes from the end of one cell to the beginning of another (.5)
and it starts to umm the whole thing again
until it eventally goes to the brain. (.5)

Question 8: (text and illustrations)
Describe the most direct and the least direct path that neural energy can take on it's way to the brain.

SCROING QUESTION 8

Most Direct: cone/rod -> bipolar -> ganglion ->brain (4)
Least Direct: rod/cone ->horizontal -> bipolar ->amacrine->ganglion (iterative)->brain (7)

Example of a Response From a Participant in the Controlled Access Condition

[Question 8 Score = 11]
Umm ok I guess the most direct would be the receptor to your bipolar cell to your ganglions
and then to the to the brain. (4)
And I guess the least direct way I guess would be from the receptor to through a horizontal cell to another receptor (error)
and it does that for a while
and then it might go to your bipolar cell
where it goes through a few amacrine cells.
And then through a ganglion
where that it might do and pass through another few amacrine cells

and then eventually to the brain. (7)

Questions That Require Generation of New Information

Question 4:
Explain as many plausible structural differences between the human visual system and that of a nocturnal animal's that would allow them to see better in the dark.

Scoring for Question 4
.5 for description of each plausible structural difference & .5 for explanation of consequence of such difference on ability to see (unbounded score)

Example of a Response From a Participant in the Controlled Access Condition

[Question 4 Score = 1.5]
Ok well ah first of all they probably have umm pupils that can dialate really big (.5)
so they let alot more light in. [function] (.5)
And then ah in terms of ah the rods and cones maybe they have ah like more cones around the periphery
and so if they get some light
it's going to go directly to ah directly through rather than
as opposed to into one from a whole bunch.
Or maybe they have ah instead of cones focused in the center and rods around the thing
they have like an equal mixture all around. (.5)
Or something like that.

Question 6:
Describe as many reason as you can for why someone might need glasses. Explain how glasses help.

Scoring for Question 6
.5 point each for description of a structural problem and .5 for an explanation of the consequences &1 point for what glasses do (i.e., refract light)
(Total score unbounded)

Example of a Response From a Participant in the Free Access Condition

[Question 6 Score = 4.5]

Oh ok.

Umm umm I don't know which one is which

but there is nearsightedness and farsightedness.

Also known as hyperopia or myopia.

And ah there's also astigmatism.

And ah the reasons for farsightedness or ah nearsightedness are due to umm improper focusing ah focusing ability.

Our ah contraction ability of the ciliary muscles they're either too strong

and of course the lens to contract (.5)

and that's what causes the ah the focal point of the eye to shift. (.5)

That () by the time the light rays hit the retina they're either too big or too small.

Umm and that's () the focal point would shift one way or the other depending on whether it's nearsightedness or farsightedness or depending on whether the ciliary muscles are too strong or too weak. (1)

And so that of course causes () to go either one way or the other.

E: And what would glasses do?

S: Glasses are either concave or convex lenses placed at an appropriate distance from the eye. (1)

That will either refract or umm ya will refract the light rays either ah one way or to one extreme or to the other.

So that will be properly refraced by the time it hits the lens.

Umm and then there's umm other possibilites.

Ahh not not sclerosis

but umm that's something else

but umm just sets in with age

where the where the mucles are too weak (1)

and of course it causes all sorts of ... E: You also mentioned astigmatism?

S: Ya.

I know that's.. the way a doctor described to me once was ah if you were for example to see a T

umm the the vertical part would not be qutie in line with the horizontal part.

and I know there's all sorts of different ah umm visualizations using prisms and that sort of thing.

Umm and I think what it is is that once something is transmitted I don't know
horizontally or something (.5)
then it gets focused to a point.
Something like that.
I can't even remember what it is
ya I'm not exactly sure what an astigmatism is.
Or how how to correct that.
But ah but ah ya the name escapes me as to as to old age.

Appendix J

Example of Coded On-line Protocol

Interp	Location	Seg	Protocol	Information Type	Process Code
1	S 6	1	So I'm wondering		metacognition
1	S 6	2	what's the other 30 percent?		metacognition
2	S 11	3	So that's the pigmented part	Structure	interpretation
3	S 17	4	Ok I'll just make an interpretation about the thing before I...	Function	
3	S 17	5	I thought when the pupil contracted	Function	prior knowledge
3	S 17	6	that it was just all the muscles in your eye relaxed	Function	prior knowledge
3	S 17	7	and then here it just said that that there is muscles which contract to make the pupil dilate.	Function	interpretation
3	S 17	8	So it's just....		
4	S 22	9	Ok so now I understand		metacognition
4	S 22	10	I always thought that there's just the lens	Structure	prior knowledge
4	S 22	11	and that the cornea and the lens were the same thing.	Structure	prior knowledge
4	S 22	12	But now I realize that it's the lens actually does the the rest of the work	Function	interpretation
4	S 22	13	I thought it was all the cornea or all the lens	Function	prior knowledge
4	S 22	14	cause I thought it was the same thing.	Structure	prior knowledge
4	S 22	15	Ok then () I'm actually learning something.		
5	S 36	16	Ok I'm remembering a science project about the blind spot in the eye		prior knowledge
5	S 36	17	you close one eye		prior knowledge
5	S 36	18	and like then you stare at something		prior knowledge
5	S 36	19	and (until) you couldn't see it.		prior knowledge
5	S 36	20	And it was all but ah because the connectors are there.	Structure	prior knowledge
6	S 45	21	So I'm wondering why so many more rods?	Structure	metacognition
7	S 46	22	I was thinking that the cones would be should be more important	Function	metacognition
7	S 46	23	so I was thinking there should be more cones.	Structure	metacognition
8	S 48	24	Ok that's why.		metacognition

Interp	Location	Seg	Protocol	Information Type	Process Code
9	Ill 2	25	This drawing is pretty complicated.		comment
9	Ill 2	26	If they just had like one part of it instead of like		
9	Ill 2	27	it would be simple.		
10	S 64	28	So I'm wondering why are they transmitting the signals across the eye?	Energy	metacognition
10	S 64	29	What's the point?	Function	metacognition
10	S 64	30	I don't know.		metacognition

Note. Interp # = the on-line interpretation episode; Location = the source location from which the on-line interpretation episode was initiated; Seg = segment number.

Appendix K

On-line Processing Event Code Definitions and Examples

Code	Definition	Example
Interpretation	Participant interprets or paraphrases literal content of text or illustration.	sf6: And basically the front of the eye is called the cornea and that's responsible for.. the shape of the cornea is responsible for like 70% of the focusing. (s6)
Prior Knowledge	Participant recalls prior knowledge and may or may not directly relate it to the text or illustration content.	sti6: I'm just trying to compare umm that umm that last thing on the screen with ah Einstein's photoelectric effect just to find to find a parallel in physical sciences. (s67)
Text Integration	Participant connects disparate text units (i.e., across sentences, paragraphs)	sf2: So I was just umm trying to ah just sum up the meaning of the last three sentences () the ciliary muscles tighten equals increased curvature equals focusing power.
Media Integration	Participant connects information across text and illustrations.	sf3: I was looking for umm the ciliary muscles but um that are mentioned in the text (follows p1 access after reading s25)
Metacognitive	participant monitors and evaluates their comprehension of the material by evaluating the comprehensibility or familiarity of the material, confirming their understanding, stating a question, expectation or prediction, identifying a comprehension problem.	Spo6: I don't know how it's ... or where the ... or what kind of stimulus the amacrine cell is carrying. (P24)
Strategy	Participant describes a particular comprehension strategy they are using or typically use when reading for understanding. Examples include rereading, visualizing, using pk to guess or reason, memorizing, skimming, ordering information,	sto1: OK I'd just like to visualize this so I'm going to go back to the previous sentences and sort of build like a model in my head. Where these specific things are. (s13)

Code	Definition	Example
Strategy (continued)		sf2: So just there I was just umm umm just trying to summarize or just trying to sum up like what each one does there. So what make careful note of I just kind of slowed down a bit there So I could make sure I could differentiate which one is which. (s45)
Comment	Participant comments on the text, illustration, or task - not content relevant.	spo3: Well the arrows definitely help. (p24) sti3: I was wondering why they told us the point before umm because it doesn't seem to lead anywhere.(s57) sti4: Um well that was an interestining line to read (s6)

Appendix L

Coded Example of Verbal Recall Protocol

Seg	Protocol	#Props	Prop Match	Info Type
1	Basically I understood that ah light goes in the eye through the cornea	3	7; 7.1; 7.2	Energy
2	and then through the pupil the black spot on your eye	3	13.2; 13.3; 13.4	Energy &Structure
3	and ah ah oh shoot			
4	I forget the coloured part of the eye.	1	11.1	Structure
5	It doesn't matter			
6	anyway			
7	And then behind the pupil is the lens.	1	18	Structure
8	And the lens focuses the light.	1	22.7	Function
9	But it only focuses like 30% that the cornea hasn't focused.	2	22.7; 6.2	Function
10	And that's (the part we can) control really	1	24.1	Structure & Function
11	that's how you you focus your eyes like the lens part of the eye.	3	24.1;25; 25.1	Structure &Function
12	And then behind that is the ah I think it's called the viscous humerous.	2	29.1; 29.3	Structure
13	And that has is a liquid that's got the same umm refraction refractive index as ah the lens	2	28.3; 28.4	Structure &Function
14	so so (there's no more) refraction	1	28.5	
15	the light just follows it's path.	1	28.6	Energy
16	And then it goes to I can't remember that spot that you have that you see the best with it	3	30; 34; 34.2	Energy &Function
17	() one degree there.		not enough	Structure
18	Ya			
19	it begins with an f.			Structure
20	but I can't remember what it is			
21	it doesn't matter.			
22	Ah ya and you get almost all your cones are there	2	46; 46.1	Structure

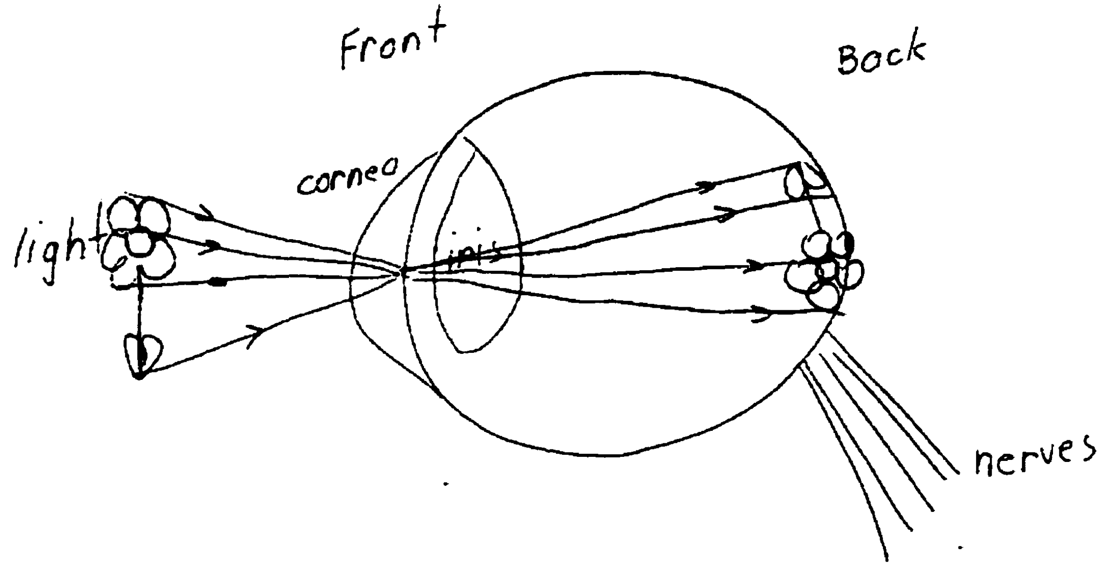
Seg	Protocol	#Props	Prop Match	Info Type
23	and the single..and the cones are the ones that that are will precise	2	44(.5) 44.1(.5); 44.2	Function
24	that that that you can.			
25	there's the cones			Structure
26	and there's the other ones that I can't remember what there called.			Structure
27	Ya			
28	The stick ones.			Structure
29	Those are the receptors those are the photoreceptors.	3	40; 40.3; 40.4	Structure
30	And they they convert photons into (ah) an impulse that your brain can understand.	3	37; 37.1; 37.2	Function
31	It goes through (like) the ganglion and...	1.5	51; 51.2 (.5)	Energy
32	but I wasn't really interested in that.			
33	and then from the ganlion all the little specific things that () the message to the brain	3	51.1; 65; 65.1	Energy
34	() I didn't think it was interesting			
35	but boring			
36	it was more just the general type of ya know			
37	Ya I like the specific...			
38	I really I really was more interested like the specific part of the eye where you've got one			
39	that everything else out of that is just is just ah....			
40	you know you detect more movement or stuff			
41	you you can't really ()			
42	your eyes have to ah ()			
43	because of the thing in the back			Structure
44	() I can't remember the spot			Structure
45	and ah basically (that's it)			

Seg	Protocol	#Props	Prop Match	Info Type
46	The iris the iris is the part of the eye with the the pigment	2	11; 11.1	Structure
47	but I don't remember ()			
48	but ah ya that's it.			
49	That's about it.			

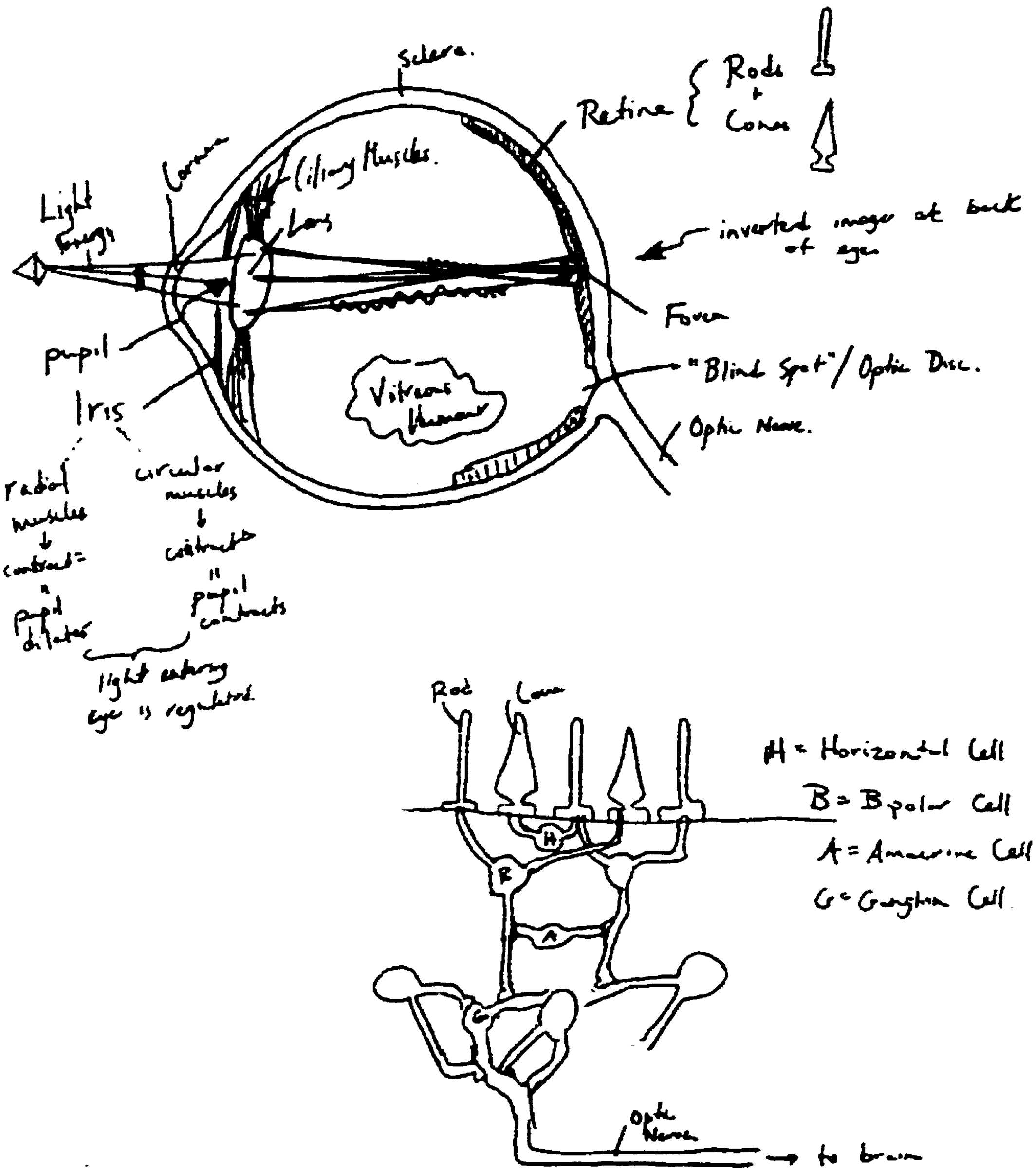
Note. Seg = segment number of line in protocol; #Prop = number of propositions recalled; PropMatch = specific propositions recalled; Info Type = type of information recalled.

Appendix M

Examples of Visual Recall Protocols



Example of Visual Recall Protocol for Participant in the Illustrations Only Condition



Example of Visual Recall Protocol for Participant in the Free Access Condition

Appendix N

Recall Organization Examples

Illustrations Only

This participant demonstrated a temporally marked narrative organization in the openings of their verbal recall protocols, introducing structural information in the context of energy information. Five of the six participants in this condition demonstrated similar opennings.

1. Umm the when the image comes the... when light comes ah [energy]
2. it first ah goes on the the retina [energy]
3. and then there is a muscle iris [energy]
4. and ah it passes through the pupil [energy]
5. then the lens. [energy]
6. And goes through the vitreous humor to the fovea [energy]
7. where it converges. [energy]
8. And it's read by the optic nerves [function]
9. And then it goes to..then there's a rod cell [structure]

Free Access

This individual demonstrates a dual organization sheme around both structure and function information. New information is introduced by describing structural and then functional information. Four of the six participants in this group introduced information through structural descriptions, one used functional organization, and one used energy information.

1 The eye itself is is umm made up of many different components [structure]
2 that have ah specific functions. [function]
3 And umm starting outwards you have cells [structure]
4 like you have a layer called the cornea [structure]
5 Which is umm transparent
6 and it performs about 70% of the of the refraction of light [function]
7 that's bending the light into into the eye. [energy]

Controlled Access

This individual demonstrates clear structure organization of information. Four of the six individuals in this group demonstrated similar openings with the remaining two individuals using energy information as a way of organizing their openings.

1 First of all the eye is a sphere
2 with a hole in the ah in the front (there) well. [structure]
3 It's covered with a transparent membrane called the sclera. [structure]
4 Umm the cornea is ah is a transparent umm organ
5 that's in front of the pupil [structure]
6 that's the hole.

Text Only

This individual also demonstrates clear structure organization of information. Four of the individuals in this group demonstrated similar openings and two used energy descriptions to organize their introductions.

1 I'm just going to picture this in my head.
2 The sclera is the protective layer covering the eyeball. [structure]
3 Umm then there's the cornea. [structure]
4 And the pupil () the iris OK the iris which is the muscular fiber. [structure]
5 And it's made up of radial and circular muscular fibers that ah expand and contract umm around the pupil to increase or decrease t to let the amount of light in. [structure & function]