Memory in Comprehension and Problem Solving:
A Long-Term Working Memory

K. Anders Ericsson & Walter Kintsch
Institute of Cognitive Science
University of Colorado
Boulder, Colorado 80309-0344
303/492-1492

Institute of Cognitive Science
Publication Number 91-13

Abstract

To account for the large demands of working memory during text comprehension and expert problem solving it is proposed that the traditional models of working memory involving temporary storage have to be extended to include a long-term working-memory portion. According to the proposed theoretical framework cognitive processes are viewed as a sequence of stable states representing end products of processing. In skilled activities these end products are stored in long-term memory and kept directly accessible by retrieval cues in short-term memory as proposed by skilled memory theory. These theoretical claims are supported by a review of evidence on memory in text comprehension, problem solving, decision making and diagnosis.

A rich literature on memory has accumulated over the more than 100 years since researchers began to study this phenomenon in the laboratory. Many studies offer models and theories to account for the major role that memory plays in complex cognitive activities such as comprehension and problem solving. In this article, we consider whether what is known from laboratory studies adequately explains the role of memory in these activities. We conclude that it does, but with an important proviso.

A limited-capacity working memory and short-term memory buffer are central constructs in the memory theory that has emerged from laboratory study. Since Ebbinghaus introduced the nonsense syllable, however, laboratory study has concentrated on situations that minimize the role of meaning and knowledge. Similarly, the classical work on problem solving has focused on unfamiliar problem domains in which the effect of subjects' previous knowledge could be neglected (Newell & Simon, 1972). Even studies of text comprehension have not dealt explicitly with the role of knowledge in comprehension (Kintsch & van Dijk, 1978). We conclude that current memory theory provides a good explanation of memory in comprehension and problem solving in studies where knowledge effects were intentionally minimized. But because this theory describes only the limiting, zero-knowledge case, it is seriously incomplete, although not incorrect.

In recent years, researchers have studied problem solving and decision making with experienced subjects in knowledge-rich domains. Comprehension has come to be viewed as an interactive process during which information from the text is combined with information the comprehender knew beforehand. At the same time, the laboratory study of memory has expanded to include research on memory in expert performance and produced the skilled-memory theory (Chase & Ericsson, 1982; Ericsson & Staszewski, 1989). The new studies of complex cognitive processes require a new view of the role of memory. In particular, as we show in this article, working memory, when subjects are skilled and the situation invites the use of their knowledge, may expand to include directly retrievable portions of long-term memory - what we have called the long-term working memory.

Historical Background

At the time of Ebbinghaus's (1964/1885) pioneering research on memory, it was generally agreed that individuals' experiences and knowledge most dramatically influenced the speed and facility with which information is stored in memory. Given that individuals' previous experiences and acquired knowledge differ greatly, it seemed difficult, if not impossible, to identify general laws describing the storage and retention of information. To address this difficulty, Ebbinghaus identified a large set of stimuli (nonsense syllables), for which no prior associations would be available. If sequences of nonsense syllables were presented at a rapid rate, Ebbinghaus believed that the influence of prior knowledge on his subjects' memory could be essentially eliminated. Hence the memorization of lists of nonsense syllables was assumed to reflect basic memory processes that encode stable associations between the individual syllables. With a fixed presentation rate, the primary factor influencing storage was the number of exposures of the items on the list. Ever since Ebbinghaus's studies of himself as the only subject, the general findings regarding speed of storage and forgetting have been replicated many times and extended to a wide range of more familiar types of materials, such as lists of digits and words. This line of research has consistently shown that human memory for arbitrary lists of items is poor and that the time to reliably store new patterns (chunks) in long-term memory is remarkably long - estimated at about 10 seconds per chunk (Simon, 1974).

Research on memory for lists of arbitrary items also showed that subjects could recall shorter lists and the last items of longer lists very accurately and with brief presentation rates. This type of recall did not reflect information stored in long-term memory and corresponded to actively maintained information in a temporary store (short-term memory). In his classic paper, Miller (1956) showed that short-term memory performance was invariant across a wide number of stimulus domains, such as consonants, digits, and unrelated words. The number of recognized patterns (chunks) that subjects could recall was found to be seven plus or minus two. Atkinson and Shiffrin (1968) proposed a model with short-term and long-term memory storage to account for a number of results from memory experiments with lists of arbitrary items. According to their model, subjects presented with items rehearsed them in short-term memory for a time sufficient for encoding them in long-term memory.

Newell and Simon (1972) showed that more complex cognitive activities such as problem solving in unfamiliar domains also seemed to be constrained by the amount of information that could be kept in short-term memory. They agreed that the cognitive processes could be described as a sequence of states characterized by the information available in short-term memory. A given state of short-term memory provided the necessary input to cognitive processes that delivered their results to short-term memory, thus altering the contents of short-term memory and thus generating a new state.

Newell and Simon's (1972) theoretical framework points to a critical difference between the information that can be held in the limited capacity of short-term memory and the information that can be retained in the vast capacity of long-term memory. The information in short-term memory is readily available for further processing; information in long-term memory needs to be accessed by a retrieval process that distinguishes specific information from the vast amount of other information stored in long-term memory. The belief that storage in long-term memory requires considerable time and that reliable access from long-term memory is difficult to attain has led investigators to discard the possibility that long-term memory can be used as a form of working memory. Laboratory studies showed that a wide range of cognitive activities such as problem solving (Atwood, Masson, & Polson, 1980; Newell & Simon, 1972), concept formation (Levine, 1966), and decision making (Payne, 1976) could be successfully accounted for by models relying on storage of intermediate products only in short-term memory.

In parallel with the study of slow, deliberate memorization of lists of unrelated items that were stored in long-term memory, investigators started to examine incidental memory of stimuli in more meaningful tasks, such as ratings, judgments, and evaluations. The most intriguing result was that incidental memory in some judgment tasks was as good as the memory of subjects who deliberately memorized the same stimuli (Craik & Lockhart, 1972; Hyde & Jenkins, 1973). Furthermore, memory of judged stimuli did not benefit from subjects' knowing about a subsequent recall test in advance. Hence memory of the stimuli must be a direct consequence of the cognitive process generating the judgment.

The classical model of short-term memory was also challenged by the vastly superior memory performance of experts compared to that of other adults. Chase and Simon (1973) showed that even after brief exposures, chess masters were able to recall most of the locations of chess pieces in meaningful chess positions. In their original theoretical account, Chase and Simon (1973) proposed that the superior memory of chess experts was mediated by a limited number of chunks of familiar configurations of chess pieces stored in short-term memory. However, Charness (1976) and Engle and Buchstel (1976) found that information about chess positions was indeed stored in long-term memory. These researchers demonstrated that other tasks interpolated with memorization to eliminate the influence of information stored in short-term memory had no or minimal effect on recall performance. Subsequent research has shown that superior incidental memory for a chess position can be obtained by having chess experts simply select the best move for the position (Charness, 1981a; de Groot, 1946/1978; Lane & Robertson, 1978).

Other research has shown long-term memory after brief exposure (less than 1 second) to meaningful stimuli such as pictures (Potter & Levy, 1969). And in a wide range of complex cognitive activities such as text comprehension (Kintsch, 1974) and decision making, substantial incidental memory for meaningful stimuli is well documented. All of this research clearly demonstrates rapid storage in long-term memory during meaningful cognitive processing. It must therefore be possible for individuals to use long-term memory for storage during complex cognitive processing. However, although many investigators acknowledge storage of information in long-term memory, they reject the possibility that such information can be retrieved with sufficient speed and reliability to meet the necessary demands of an extended working memory.

Skilled-Memory Theory

After studying the effect of extended training on memory performance in tasks such as digit span that are assumed to measure short-term memory, Chase and Ericsson (1982; Ericsson, 1985; 1988) proposed skilled-memory theory. This theory explains how subjects, after training, can use long-term memory as an extended working memory with storage and retrieval characteristics similar to those of short-term memory.

To achieve rapid storage in long-term memory, skilled-memory theory claims that an individual must encode the presented information by relying on associations with pre-existing knowledge and patterns in long-term memory. This claim is consistent with the research reviewed earlier showing substantial incidental long-term memory for information processed during meaningful cognitive activities such as reading a text. Chase and Ericsson (1982) extended this claim by showing that with sufficient practice subjects could acquire meaningful encoding methods for initially meaningless information such as lists of digits. Several of their subjects were collegiate runners who used their knowledge about running times to encode groups of three and four digits. For example, 3493 could be encoded as the world-class running time for the mile, that is, 3 minutes 49.3 seconds. To uniquely encode all digits, especially the last digit or digits, the subjects relied on additional associations involving numerical patterns, such as the fact that the last digit in 3493 is the same as the first. The encoded digit groups were genuinely stored in long-term memory and were not simply retrieved, pre-existing chunks: The subjects were able to recall virtually all of the digit groups from all the presented lists of digits during post-session recall, amounting to between 200 and 300 digits. After reviewing other studies of memory experts and individuals with alleged exceptional memory, Ericsson

(1985; 1988) argued for the generality of the claim that individuals can acquire methods for meaningful encoding of meaningless materials such as lists of random digits and unrelated words.

If individuals can use long-term memory as an extended working memory, the critical issue is how they can efficiently retrieve information once it is successfully stored in long-term memory. Unlike the limited information kept in attention or short-term memory, information stored in long-term memory must be retrieved with sufficiently specific retrieval cues before it can be processed again. According to skilled memory theory, individuals overcome this problem by associating the encoded information with special retrieval cues belonging to a retrieval structure at the time of the original presentation. Whenever individuals need to recall the stored information, they activate the special retrieval cues to retrieve the encoded information from long-term memory. The theoretical notion of a retrieval structure is very general and is instantiated in different ways to accommodate different demands on subsequent retrieval.

Chase and Ericsson (1982) focused their empirical and theoretical analyses on serial recall of digits. Their trained subjects were able to recall lists of over 80 rapidly presented digits by segmenting the digits into groups according to a predetermined scheme. They then encoded each digit group and associated it with a unique location in their hierarchical retrieval structure as illustrated in Figure 1.

Insert Figure 1 about here

At the time of recall, the subjects can easily regenerate any of the unique locations of the retrieval structure and use a given location as a cue to retrieve the corresponding digit group. By proceeding sequentially through the retrieval structure the subjects can serially recall all of the digits in their presented order. According to skilled memory theory the encoded digits groups are not stored in short-term memory but are directly accessible based on retrieval cues stored in attention/short-term memory. Chase and Ericsson (1981) demonstrated this form of accessibility experimentally by using a cued recall task. After the subject had memorized a digit sequence they presented locations in the retrieval structure and asked the subject to recall the corresponding digit group or asked the subject to point to the location of a presented digit group. Especially with extensive practice on the digit-span task such retrieval is virtually immediate (Staszewski, 1988b). Further evidence for this immediate and flexible retrieval using retrieval structures is shown by their ability to memorize matrices of digits without any specific training, and in particular to retrieve these digits in many different recall orders, as illustrated in Figure 2.

Insert Figure 2 about here

The trained subjects' storage and flexible recall matched those of the exceptional subjects studied by Binet (1894) and Luria (1968) and would thus meet the criteria specified by these investigators for exceptional visual/photographic memory. Other studies of memory experts and subjects with alleged exceptional memory recall (reviewed by Ericsson 1985, 1988a) almost always yield evidence for the explicit use of retrieval structures in exceptional serial recall. Often the retrieval structures are hierarchies of spatial locations, but frequently subjects rely on a sequence of locations as in the method of loci (Yates, 1966).

The demands for retrieval in serial recall of presented lists differ greatly from retrieval demands in complex cognitive activities. During complex mental calculation, individuals must store the numbers of the original problem as well as intermediate products. Staszewski (1988a) showed how subjects achieving exceptional performance in mental multiplication develop retrieval structures to store and subsequently retrieve intermediate products as an integral part of their skill. The skill of playing blindfold chess, in which a subject has to search a mental representation of the current chess position, was found by Ericsson and Oliver (Ericsson & Staszewski, 1989) to be mediated by retrieval structures. After a brief presentation of a chess position the chess master could immediately recall which chess piece was located in a given square of board when the corresponding location was presented as a cue. The analysis of memory for chess positions also revealed efficient access of semantic relations between chess pieces and configurations of chess pieces.

Even in the domain of expert memory for digits, Chase and Ericsson (1981, 1982) documented evidence for semantically driven retrieval. They found that their trained subjects, in encoding a digit group in a presented list, would automatically retrieve other digit groups presented earlier in the session and encoded with a similar running-time encoding. Attempts to measure the speed of this form of semantically based retrieval suggest access in less than half a second. By comparing reaction times in a recognition task for cases involving retrieval from short-term memory with cases involving retrieval from long-term memory Anderson (1990, Table 6-2) estimated the retrieval from long-term memory to be 420 ms for briefly studied items and 280 ms for more extensively studied items. The retrieval time for the component elements of chunks was reviewed by Yu, Zhang, Jing, Peng, Zhang and Simon (1985) and was estimated to range from 281 ms to 373 ms. Staszewski (1988b) estimated that a highly trained memory expert retrieved chunks from long-term memory in around half a second. Fast and essential;ly immediate retrieval from long-term memory requres appropriate retrieval cues in attention/short-term memory. If the retrieval cues in short-term memory are lost due to a demanding interpolated task, then the retrieval cues have to be generated before the desired information can be accessed from long-term memory with much longer retrieval times as a consequence. In his study of recall of chess positions Charness (1976) found that the latency of recalling the first chess piece increased by 2-4 s when a 30 s activity was interpolated between presentation and recall. When Ericsson and Oliver (Ericsson & Staszewsky, 1989) studies cues recall of chess pieces from two memorized chess positions they found that retrieval times were almost 1 s slower when the cued locations alternated betweeen the two positions as compared to when the cued locations referred to the same position. Hence, consecutive retrievals from the same position can be done with the same retrieval cues in short-term memory but a change in the probed chess position requires a change of the corresponding retrieval cues in short-term memory. In sum, there are many possible methods of retrieving encoded information from long-term memory if, at the initial encoding and storage of information, an individual can anticipate future contexts for retrieval.

The skilled-memory account of extended working memory in skilled complex activities is attractive because it is completely consistent with the limited capacity limitation of attention, short-term memory, and sensory stores proved by laboratory studies. In fact, detailed analyses of memory experts' encoding processes show that the number of independent items encoded into a single group is virtually always 3 or 4, and larger groups of items are nearly always broken down into hierarchically joined subgroups (Ericsson, 1985).

Furthermore, skilled-memory theory is based on current theories of storage in cue-based retrieval from long-term memory. According to skilled memory theory, extended working memory is limited to those activities in which subjects have sufficient knowledge for efficient encoding and storage in long-term memory and have acquired retrieval structures and procedures for reliable and fast retrieval of desired information. Meaningful encoding of a stimulus implies that an individual selects one of several possible meanings or aspects. Hence, unlike storage in short-term memory where presented stimuli can be stored without a semantic commitment, storage in long-term memory implies a semantic disambiguation, making a subsequent reinterpretation of the meaning impossible unless the ambiguity and the alternative interpretations were explicitly encoded. Similarly, the encoding of information and its association with a retrieval structure imposes clear limits on the way the information can be retrieved. At the time of encoding, considerable skill is required to determine in what contexts the information is relevant and should be efficiently retrieved. Under these limiting conditions, subjects can use long-term memory as an extension of their working memory.

In this paper we review the role of memory in complex cognitive activities. We first present a general theoretical framework for the structure of complex cognitive processes and their dependance on various types of memory. Within this framework we propose some general hypotheses about the structure of the memory system and explicate general empirical predictions. Then we use this theoretical framework to review findings from the literature on text comprehension, problem solving, decision making, and expert performance.

Toward a General Framework for the Nature of Memory in Complex Cognitive Activities

All cognitive processes require time, and complex cognitive processes often require substantial time to run their course to completion. Ever since Aristotle, complex cognitive activities such as thinking have been described as a sequence of thoughts. As a first approximation we can therefore describe complex cognitive processes as an ordered sequence of states, as illustrated in Figure 3. In this article we will refer to the temporal ordering of mental states as the horizontal dimension of thinking.

Insert Figure 3 about here

If we could accurately describe a mental state at any point in time, we would be able to summarize all the effects of the preceding cognitive processes upon those directly following. Given the vast complexity of all neural activity in the brain at any one time, it seems impossible to identify such a full description of a mental state. However, the critical assumption of cognitive psychology is that the impact of neural activity can be summarized by a limited number of generated results or products. Hence, if the temporal dimension is divided into intervals of an appropriate length, cognitive processes can be described as a sequence of generated products. Newell and Simon (1972) proposed that the contents of short-term memory were sufficient to characterize the sequence of mental states corresponding to cognitive processes. More recently, Anderson (1987a) argued for the distinction of macro and micro processes, macro processes roughly corresponding to those processes that generate stable products affecting the sequence of the thought process. In his recent model of text comprehension, Kintsch (1988) proposed that successful text comprehension can be described as a sequence of states mediated by cycles of construction and integration of segments of the text.

In Figure 3, the cognitive processes are divided into sequences of segments with completed processing leading to stable results. Within a segment, various processing occurs involving memory buffers and a wide range of interactive processing, which is illustrated by the vertical dimension in Figure 3. While a semantic interpretation of the processed information is extracted, the presented information is maintained in temporary buffers containing differing levels of activated information ranging from perceptual to conceptual.

The primary difference between this model and the traditional one is its claim that information in attention can be stored in long-term memory in retrievable form. With appropriate encoding, association with retrieval cues, or both, generated information can be rapidly retrieved with the correct retrieval cues, even when this information is no longer active in attention or short-term memory. In the more traditional model of short-term memory it has always been clear that not all the information about a complex chunk was stored directly in short-term memory. Instead, storage of a complex chunk in short-term memory meant that a single label or long-term memory address was stored or kept active in short-term memory. This label could be used to rapidly retrieve elements of the chunk from long-term memory within around 300 ms (Simon, 1979, Chapters 2.3 and 2.4). Hence the primary difference between the classical conception of working memory and the one proposed here concerns the ability to store new complex memory traces in long-term memory during cognitive processing. In our theoretical framework, working memory consists of two components, as shown in Figure 4: 1) the activated portion of long-term memory which holds the endproducts of the comprehension process in the focus of attention/short-term memory, and 2) the portion of long-term memory that is directly retrievable via retrieval cues in attention (long-term working memory). Because traditional models of working memory include only the first of these, we use the term extended working memory to distinguish our framework.

Insert Figure 4 about here

Only under restrictive conditions can long-term memory be used as an extended working memory. Namely, as Chase and Ericsson (1982) concluded, the stored information must be encoded with pre-existing knowledge and patterns available in long-term memory so that appropriate retrieval cues can readily access it. In the following sections of this article, we argue that both expert cognitive activities and skilled activities, such as reading meaningful text, meet these conditions. However, a wide range of laboratory studies of problem solving, concept formation, and decision making have used unfamiliar task domains. The empirical results from these cases show that short-term memory is sufficient for describing the mental states of cognitive processes. Under such circumstances, our theoretical framework predicts no or very limited use of long-term memory and is consistent with Newell and Simon's (1972) account for these findings.

In Figures 3 and 4 we sketched the framework for the use of memory in complex cognitive activities. We now elaborate it in the context of both text comprehension and problem solving. In our review of empirical findings on the function of memory, we first discuss text comprehension and then problem solving, concept formation, decision making, and expert performance.

The task of comprehending a text is particularly well suited to the analysis proposed here. While reading a well-written text on a familiar topic, all skilled subjects read the text in more or less the same smooth and linear fashion. The sentences and phrases constitute natural segment boundaries for processing. Even the words within a phrase are scanned in a linear, orderly fashion, a phenomenon which allows the study of the cognitive processes operating within segments. Successful comprehension of a text involves the predictable integration of information across sentences in a similar manner across all the subjects reading the text. Hence, as a first approximation, we can argue that comprehension of a text involves the same sequence of states and segments for all subjects. In direct contrast, cognitive processes and the corresponding sequences of states differ greatly between subjects in studies of problem solving, decision making, concept formation, and expert performance. Consequently, our analysis of these complex cognitive processes is necessarily organized around more general results and mechanisms.

Memory in Text Comprehension

Text comprehension is a prototypic information processing activity. During reading, the proximal stimulus (printed words) is transformed into integrated thoughts. Initially, an external stimulus is represented as an optic array impinging on the reader's retina. This information undergoes a series of transformations (Marr, 1982) before conscious perception occurs, and is further transformed in the thinking stages of information processing. As the information is transformed, new levels of representations emerge: The perceived object is represented by a name, which in turn activates additional information from long-term memory. The object becomes integrated into a representation of the situation experienced, is then elaborated by inferential processes, and so on. A reader's goal is to form in long-term memory an integrated representation of the information presented in the text. Because reading is a strictly sequential process, the reader achieves this integrated representation in an incremental fashion, drawing on intermediate representations at different levels of analysis of the text.

The psycholinguistic literature has widely used a distinction, derived from linguistics, between three aspects of language, namely, syntax, semantics, and pragmatics. Comprehension of a sentence is thought to start with a syntactic analysis, followed by a semantic interpretation, and possibly some pragmatic elaborations. It has not been possible, however, to cleanly separate these processes in comprehension. We doubt, therefore, that this is a useful approach for analyzing the psychological process of comprehension (we do not question its use in linguistic analysis). We prefer another distinction that is based more directly on the nature of the processes involved in discourse comprehension (Hörmann, 1976; van Dijk & Kintsch, 1983).

Different mental representations result from the linguistic processing of the language of a text, the understanding of the text as a separate mental object, and the integration of the textual information and background knowledge. It is useful in general to distinguish three different levels of representation.

1. Linguistic surface structure. This component of the memory representation of a text comprises the traces of the words of a text, not only by themselves but syntactically, semantically, and pragmatically interpreted in the context of a sentence. The surface structure is generally stored until the end of a sentence and may be lost rapidly thereafter.

2. Propositional textbase. The textbase is a coherent conceptual representation of a text and its structure. It contains both a microstructure and a macrostructure. Micropropositions are usually directly derived from the text (they correspond to phrases and sentences), but may be the result of inferential processes (e.g., to bridge a gap in the coherence of a text). The macrostructure results from selection and generalization processes operating on the micropropositions. Macropropositions are thus partly cued directly by the text, and partly inferred. The textbase is stored in working memory for the sentence currently being read. Thereafter it can be retrieved from long-term memory by means of conceptual retrieval cues.

3. Situation model. The model of the situation described by the text integrates textual information and background knowledge. The reader's representation is not strictly propositional; it may be a spatial model such as a

map. The situation model provides the reader with a basis for inferencing, elaboration, and action. In most cases, because the reader's purpose is to construct a new situation model or to update an already existing one, this model is usually a major and long-lasting component of the memory trace.

Researchers concerned with memory for text have focused on different levels of representation. Glanzer, Dorfman, and Kaplan (1981) claimed that it was sufficient to consider the surface level, at least as far as short-term memory was concerned. Most researchers were unwilling to accept such a restriction and differentiated at least between a surface level and some sort of semantic representation (Frederiksen, 1975; Meyer, 1975; Kintsch, 1974; Schank, 1972; Schank & Abelson, 1977). On the other hand, it has been repeatedly demonstrated that although surface representations are unavailable in many cases, subjects manage to solve experimental tasks very well on the basis of their situation model (Bransford, Barclay & Franks, 1972 Bransford & Franks, 1971; Johnson-Laird, 1983).

Experimental evidence for the division of the representation of text into three levels lies in the systematic differences in memory strength that researchers have demonstrated between these levels of representation (Schmalhofer & Glavanov, 1986; Fletcher & Chrysler, 1990). If subjects read a text and are later tested with sentences from that text as well as with distractor items, their ability to differentiate between old sentences and meaning-preserving paraphrases provides evidence of surface memory. To the extent that subjects' answers contain more false negatives for a recognition test on paraphrases of sentences which actually appeared in the text than on inferences, subjects can be said to have retained the propositional representation of the text. Finally, an increase in the false negative rate for inferences compared with contextually related but non-inferrable distractor sentences can be considered evidence for a surface level representation. Under the conditions of the experiment by Schmalhofer and Glavanov (1986), for instance, the estimated strength of the surface, textbase, and situation model traces were d' = .14, .68, and 1.96, respectively. Subjects in this experiment were reading excerpts from a textbook on the computer language LISP. They had formed a strong situation model, but retained little of the textbase and less of the surface form of the text. It must be emphasized, however, that the level of representation subjects prefer strongly depends on the nature of the text they read as well as on their reading goals.

Zimny (1987; see also Kintsch, Welsch, Schmalhofer, & Zimny, 1990), in a sentence recognition experiment like the one just described, observed differential decay rates for the different levels of representation in memory.

Under her experimental conditions, surface memory decayed within 20 min. The strength of the textbase gradually decreased over a 4-day period, without ever quite reaching the zero level. The situation model, on the other hand, did not decrease at all during this time period.

Although these studies show the usefulness of distinguishing between various levels of representation in text comprehension, they give no justification for reifying these levels. For one study the surface properties of a text may be most interesting; for another, the textbase or the situation model. However, there is no such thing as a surface buffer, nor are there three separate levels of cognition. These are only analytic categories, perspectives we can bring to bear in analyzing texts. In one case we want to emphasize the surface relations among the text units; in another, the semantic or situation model relations. Furthermore, these three levels are not exhaustive. For example, in the study of algebra word problems, it is important to distinguish the algebraic problem model from the naive understanding of a situation (Nathan, Kintsch, & Young, in press); In human-computer interaction, the distinction between a system model (a task analysis from the standpoint of the system) and the situation model (how the task is understood by the naive user) has proved to be very useful (Fischer, Henninger & Redmiles, 1991); In analyzing poetic language, the emphasis may be on rhyme, rhythm, and alliterative relations (Kintsch, in press a). Thus, the higher levels of representation in text comprehension discussed here are to be understood simply as useful categories for the analysis of comprehension processes.

To comprehend text, a reader must construct an integrated, meaningful representation, focusing on various aspects of the text from the surface structure to the situational model, which must be stored in persistent form in long-term memory. This representation is formed during the reading of the text. Following our model depicted in Figure 3, we argue that the construction of this representation can be viewed as a sequence of states which roughly correspond to the completed processing of a segment of the text, that is, phrases or sentences. Within our theoretical framework, we distinguish between the processing and integration of the current text segment, for example, sentence or phrase, and the stable semantic representation of this segment emerging as a result of this processing. During the processing of the segment, transitory representations allow subjects access to multiple meanings of words in the segment, whereas once the semantic processing of the text segment is completed, there is no simple way to change the semantic interpretation. We first discuss the representations available during the processing of text segment and the processes with which a stable semantic representation are extracted. We then turn to the information carried over from processing of one text segment to the next.

Text Comprehension: Processing Within Text Segments

Reading is by its very nature sequential: Words follow words, and chapters follow chapters. Memory must play a major role in such a process because each word, sentence, paragraph, or chapter cannot be understood by itself, but must be related to what was already known. Indeed, among the factors known to affect comprehension, those related to memory processes play a large role. A variety of memory-dependent skills are necessary for comprehension, ranging from decoding skills (e.g., Cunningham, Stanovich, & Wilson, 1990) to lexical access (e.g., Jackson & McClelland, 1979; Baddeley, Logie, Nimmo-Smith, & Brereton, 1985), to vocabulary (e.g., Baddeley et al., 1985), and at a higher level, general inferencing skills (e.g., Garnham, Oakhill, & Johnson-Laird, 1982; Oakhill, 1982; Oakhill & Garnham, 1988). However, for our purposes we need to distinguish these factors from memory for the information about the currently read text. We first discuss the transient memory representations of the currently processed text segment.

Memory Representations of the Current Text Segment

As a reader reads a segment of a text, the reader's eyes essentially fix upon each content word of the segment in serial fashion. During each fixation different types of memory representations are activated and generated. Not all representations are equal. Intermediate levels of representation are in general associated with temporary buffers and characterized by limited retrieval conditions, whereas the products of information processing, which may reach the level of consciousness, are more durable and more broadly available. This is what we have termed the vertical dimension in the role of memory in cognition. Extensive research has resulted in a high degree of consensus on the different types of representations generated during reading. Visual information is first registered by light receptors on the retina; then the neural information is further analyzed en route to the associative cortex in a series of transformations. Higher levels of representations are derived from the lower level representations as the information at each lower level is maintained briefly in associated memory buffers. We thus describe the vertical dimension of memory shown in Figure 3 as a sequence of temporary memory buffers with different characteristics. According to Potter (1983), whose proposal we have somewhat modified here, reading a text involves the following sequence of memory buffers with their corresponding levels of representations:

1. Retinotopic icon. The perceptible continuation of a single presented word resulting from photoreceptors and other neural mechanisms. Apparently, this buffer has no function in normal reading and plays a role only under special laboratory conditions.

2. Spatiotopic visual memory. Retinal information from successive eye-fixations is integrated at this level of processing. The printed text is represented as a stable structure located in space.

3. Reatopic visual memory. Spatial characteristics of the retinal information are less relevant at this level of transformation than are the configuration of visual features and patterns of the perceptually available text segment. Even when this text segment is removed and replaced with an irrelevant stimulus (visual mask), information about the original text segment is retained in this memory buffer for several seconds.

The neurophysiological bases of memory Buffers 1-3 are relatively well understood, including the differentiation between space- and pattern-information. The following memory buffers with higher levels of representation involve learned recodings of visual stimuli:

4. Acoustic short-term memory. Even skilled, adult readers transform visual information in reading into an acoustic form: Acoustic coding is preferred in short-term memory. (We ignore here the debate whether this level of representation is best understood as acoustic or articulatory-motoric).

5. Conceptual short-term memory. At this level of representation, words and objects are perceived and understood. Experimental results show that words can be momentarily understood, but are then lost because of interference from other cognitive processes.

6. Working memory. We understand working memory as the entire cognitive workspace (here we deviate from Potter, 1983) where information about the previously read text is stored in rapidly accessible form. Information in working memory may be accessed in two ways. First, a limited amount of the most recent information may be accessed via temporal context cues. This corresponds to the traditional short-term memory buffer. In addition, however, other, not necessarily recent, information may be accessed if a reader has formed suitable retrieval structures during comprehension of earlier portions of a text. Thus, for unskilled readers reading in an unfamiliar domain, working memory reduces to the current input plus whatever is still maintained in the short-term memory buffer (i.e., retrievable via purely temporal cues). On the other hand, for skilled readers reading in a familiar domain, much of the previously processed text may be rapidly accessible via retrieval structures such as those that characterize skilled memory. This corresponds to our concept of expanded working memory, in which a portion

of the reader's long-term memory structure becomes as rapidly accessible during reading as information held in the classical short-term memory store.

Thus, we distinguish between information stored temporarily in working memory and information stored in long-term memory but directly accessible through retrieval cues stored in working memory. Furthermore, some information will be stored in long-term memory but can be retrieved only after active efforts involving memory search.

Figure 5 illustrates the temporal succession of buffers after Potter (1983). The time necessary to encode information acoustically can be estimated as about 175 ms: A reader needs at least 400 ms to pronounce a word, 225 of which appear to be needed for articulation (if a subject knows which word is to be pronounced, the time to articulate it is 225 ms). Conceptual processing (Buffer 5) begins about 100 ms after the fixation of a word and requires around 250 ms. As soon as a word has been identified conceptually, Buffer 6, working memory, receives input from Buffer 5. Thus, it takes about 350 ms from fixation before a word is available for further cognitive processing in working memory. The stages of representation that are directly relevant to further cognitive processing are Buffers 4 and 6.

Insert Figure 5 about here

The sequence of buffers outlined in Figure 5 is critical to the integration and comprehension of segments of text (cf. Figure 3), such as phrases and sentences. Deficits in any one of these storage buffers may have complex results, beyond the buffer in question. Thus, phonological storage deficits affect not only the acoustic buffer itself but also general learning and comprehension capacities. Baddeley, Papagno, and Valla (1988) observed a patient with such a deficit and found a normal ability to associate meaningful, familiar word pairs, but an almost total inability to associate words with unfamiliar sounds (words in another language). Apparently, the patient could process familiar sound patterns adequately, without having to maintain them in the acoustic buffer. But the patient did not have enough time to analyze unfamiliar sound patterns without maintaining them at least briefly in the acoustic buffer. Similarly, Baddeley and Wilson (1988) observed another such patient who had no trouble identifying single words and comprehending brief sentences, even syntactically complex ones. However, this patient could not understand longer sentences because, although comprehension abilities were intact, the patient could not maintain speech sounds long enough in the acoustic store to complete a successful analysis. Analysis was unsuccessful, for instance, when an early word in a sentence could not be disambiguated until late in the sentence.

Figure 5 is misleading, however, because it is incomplete. It does not show the effect of knowledge on reading: Readers can read familiar materials much better than they can read unfamiliar ones. In Hoffmann's (1927) classical study, children in grades 1 to 8 read tachistoscopically presented letter strings. By a ratio of about 1 to 4, the children were much better when they were reading familiar words than when they were reading consonant strings. Significantly, this ratio was considerably higher (1 : 4.4) for the better students (approximately the top quarter, by teacher rating) than for the poorer students (1 : 3.2 for the lowest quarter).

Figure 6 shows some of Hoffman's (1927) results. Good and poor students performed comparably with consonant strings, but as the reading material became more familiar and more meaningful, the good students improved much faster than the poor students did. Individual differences between good and poor readers has been found for the speed of decoding of words, lexical access, and generation of inferences (Carver, 1990; Hunt, 1978). Knowledge and familiarity thus affect not only comprehension and memory, a finding we discuss in a later section, but even the reading process itself.

Insert Figure 6 about here

The Time Course of the Construction of Mental Representation of Text Segments

As shown in Figure 5, it takes about 350 ms for a word to be made available in working memory. At this point the word has been acoustically coded and identified semantically. There is more to text comprehension than this, however, in that interdependent relational structures representing the text as a whole must be formed at various levels of analysis, as discussed above. The time course of these processes has been investigated chiefly by means of priming experiments (Swinney, 1979; Seidenberg, Tanenhaus, Leiman, & Bienkowsky, 1982). A frequently used method is lexical decision. While listening to or reading a text, the subject is interrupted at a certain word, the priming word, by a presented letter sequence. The subject's task is to respond as quickly as possible whether or not the letter sequence (the target) is an English word. Reaction times can be compared when the priming word and the target word are unrelated and when they are associatively, semantically, or thematically related. Any decrease in the reaction time to the target word is attributed to the activation of the target word by the related priming word. By varying the temporal distance between the priming word and the presentation of the target word, experimenters can obtain a picture of the time course of this activation process.

An experimental study by Till, Mross, and Kintsch (1988) has focused on two kinds of relationships between the priming and target word: pre-existing associations and discourse-based, thematic relationships. In the first case, a homonym (e.g., "mint") was used as the priming word, and the target word was either a contextually appropriate or inappropriate associate (e.g., "candy" or "money"). The results are shown in Figure 7, in which the difference in the reaction time between unrelated and related target words is plotted as a function of the prime-target asynchrony. For the first 350 ms after the prime word has been presented, reaction times to related and unrelated target words do not differ; but after that time, reaction times for related target words are consistently faster.

These data replicate the original results of Swinney (1979) and others. They imply that context does not work like a filter that facilitates the expected or inhibits the unexpected. Rather, the construction of meaning in a discourse takes time, requiring the integration of the target word and its context. The 350 ms required for this integration process agrees well with Potter's (1983) estimate for conceptual identification: Once a word arrives in the working memory buffer, its meaning is fixed; the basic meaning construction process is over. It may be elaborated further, but changing it now requires explicit repair processes, as in some garden path sentences.[1]

Gernsbacher, Varnes, and Faust (1990) gave subjects sentences like "He dug with the spade" to read and then looked at the activation of associatively related but contextually inappropriate words, e.g. "ace" in the example given. They divided their subjects into skilled and nonskilled readers. For skilled readers their results were as expected: the memory representation of "ace" was activated 100 ms later but not 850 ms later. For nonskilled readers, however, "ace" was activated at both the short and the long delay intervals, suggesting that the process of contextual integration may be deficient in such readers.

Insert Figure 7 about here

---

[1] These conclusions are not uncontroversial. A number of priming studies (e.g., Tabossi, 1988) have reported filter-like, immediate context effects on word identification. At this point it is not clear what differences in methodology and materials are responsible for these contradictory results, nor how this would affect the explanation offered here.

The time course of activation for thematically related target words was quite different from that for local, associative relations. In the study of Till et al. (1988), given the sentence "The townspeople were surprised that all buildings had collapsed except the mint," subjects could not identify "earthquake" as a word more rapidly than thematically irrelevant control words when the words were presented 200-500 ms after the end of the sentence, even though it would have been possible to infer the sentence topic earlier. As Figure 7 shows, only after 1 s could a priming effect be observed, leading us to conclude that this much time was needed to infer the sentence topic for isolated sentences. Word meanings are formed quickly; discourse meaning, in this case, took about 1 s to be constructed.

These observations do not imply that certain types of processes are delayed until others are finished. On the contrary, it appears that all processing is performed as soon as possible (Just & Carpenter, 1980). But some processes take longer than others, and some depend on the outcome of other processes. Words are identified semantically while they are fixated during reading. In Carpenter and Just's (1981) experiments, some 80% of all content words were fixated. During reading aloud, however, although the eye precedes the voice by several words, word identification is not delayed by that much. Just and Carpenter (1980) have shown that familiar, high-frequency words are fixated for a shorter duration than difficult, low-frequency words. Since the meaning of a word thus determines at least in part the duration of fixation, meaning must be formed during the fixation.

Not all interpretation can be immediate, however, because in many cases a final interpretation depends on the sentence or discourse context. Hence readers take extra time at the end of sentences (Carpenter & Just, 1981; Just & Carpenter, 1980; Dee-Lucas, Just, Carpenter, & Daneman, 1982; Aaronson & Scarborough, 1976). This wrap-up time may be used for correction of preliminary interpretations that were found to be incorrect in the context of the sentence, and also for extra inferencing. In addition, it appears that topics of sentences presented in isolation are inferred only at the end of a sentence, even when that would have been possible earlier (Till et al., 1988; Figure 7). For sentences in a well-written text, the context aids the reader to integrate the presented text primarily during the actual reading of the sentence.

In addition to the longer reading times for last words of sentences, there are several other sources of evidence for the claim that a stable semantic interpretation is generated at the end of a text segment, as shown in Figure 3. The first, a technique to determine the units of cognitive processing, involved the presentation of clicks during comprehension of discourse. The pioneering work suggested that subjects perceived the clicks to be presented knowledge use in comprehension. We focus here on the memory aspects of that theory.

The construction-integration model provides a characterization of the memory trace of a text in working memory as well as in long-term memory. In general, this trace may contain surface elements and propositional elements, as well as situation model elements. Each element may enter into a set of surface relations with other elements, another set of semantic relations, and a third set of relations determined by the situation model. Not all of these relations may actually be computed during comprehension, however, or play a role in the final long-term memory trace of the text, where features relatively unimportant at the time of encoding tend to become deactivated.

A simple example cannot introduce all relevant aspects of this model, but consider the following one-sentence text: "Tina gathered the kindling as Lisa set up the tent." We need to consider both the words and phrases of this text as well as the concepts and propositions the words stand for, as shown in Figure 8. We also need to include in our analysis the knowledge activated by this text. To keep things simple, only a few knowledge elicitations are included in the illustrative example shown in Figure 8. We assume that "kindling" produces the association "fire"; "gather the kindling" produces "camping"; and "tent" produces both "camping" and "circus." To give the semantic content of the text more emphasis than the actual words, weights of 2 are assigned to semantic links and weights of 1 are used elsewhere in the network. Initial activation values are assumed to be 1 for all nodes.

Suppose the network shown in Figure 8 is processed in two cycles, the first corresponding to "Tina gathered the kindling" and the second to the rest of the sentence. If we integrate the first cycle network, the proposition GATHER[TINA,KINDLING] emerges as the most highly activated element. This proposition is therefore carried over in the short-term memory buffer and becomes part of what is processed in the second cycle. After the second cycle is integrated, long-term memory values are calculated. All nodes and links in the network are adjusted according to the activation values they received (see Kintsch & Welsch, in press, for details). These memory strength values are shown with each text element (but not for the links) in Figure 8.

Insert Figure 8 about here

Not surprisingly, linguistic elements turn out to be weaker than propositions - this is a simple consequence of the fact that we assigned stronger weights to the latter (see Kintsch, Welsch, Schmalhofer, & Zimny, 1990, for a rationale). More interesting is the pattern of the text propositions themselves. The propositions with the most connections, GATHER[.] and [SET-UP[.], are strongest; the objects, which produced knowledge elaborations, are stronger than the actors; and among the actors, there is a pronounced primacy effect: TINA is about twice as strong as LISA. This latter fact accounts for the observation reported in psycholinguistic studies (e.g., Gernsbacher, Hargraeves, & Beaman, 1989) that the first actor mentioned is more strongly primed than the second actor mentioned when tested after the sentence has been fully processed (see the analysis in Kintsch & Welsch, in press).

Of more interest here, however, is the fate of the knowledge elaborations. The contextually irrelevant but not totally inappropriate association "fire" remains with a weak memory strength (.10); the contextually inappropriate association "circus" is suppressed (.00 activation); and the contextually appropriate "camping" obtains a strength of .72, which is higher, in fact, than the memory strength of three of the seven actual text propositions. Thus, the model predicts for the choice of parameters made here that "camping" would be a significant component of the memory representation of the example sentence. The model has inferred a topic for the sentence, and it has eliminated the contextually irrelevant "circus."

Comprehension of Text Segments: An Overview

As information is transformed from patterns of light and sound impinging on individuals sense organs, a sequence of different forms of representation of this information is generated. Each of these forms of representation needs some temporal continuity, which is provided by a sequence of buffers corresponding to these various levels of representation. We have called this the vertical dimension of memory (see Figure 3). What happens for visual stimuli from the retina to the higher brain centers is not our concern here. However, at these levels of analysis, information processing appears fixed and specifiable in terms of brain processes. Buffers need not be inferred from psychological data but are identifiable brain structures, at least in principle. A wealth of evidence exists, both behavioral and nonbehavioral, for the separate representation of spatial and object information in vision. This changes at the level of working memory and consciousness. No separate brain structures have evolved to specialize for higher cognitive functions. At this level we are dealing with a general purpose brain and memory.

To generate an integrated representation of a text with many sentences, paragraphs, and chapters, a reader has to proceed sequentially, integrating one text segment at a time. The higher level representation of the text, such as the situational model, is constructed a piece at a time during the processing of each text segment. The sequential process of connecting each new text segment to the integrated representation in long-term memory can be described as a sequence of states, as described in Figure 3. During the processing of each new text segment, many different sources of information are integrated with a great deal of parallel processing and top-down influences from the previously constructed representation of the text, as illustrated in Kintsch's construction-integration model. While a text segment is still being processed, there is a fair amount of flexibility that allows a reader to recover from incorrect semantic disambiguation of word meanings and incorrect syntactic mappings due to the transient storage of the surface form of the text segment. As the processing of the text segment is completed, the semantic encoding is fixated in long-term memory and the temporary buffers now contain the activated information corresponding to the surface form of the next text segment.

We now turn to the horizontal dimension of memory, which is what we usually think of when we use the term "memory": not buffers that tie together different stages of information processing, but a store that maintains the old when new events are perceived. Usually that store retains the end product, or at least a fairly advanced one, of information processing - in the case of discourse, that product would be representations mostly at the level of the textbase and situation model. As the reader's thoughts move from one mental state to the next, what of the old thought is retained, and how is it retrieved when needed again?

Text Comprehension: Successive States

Reading can be described as a sequence of states (see Figure 3) where each state corresponds to the completed processing of a text segment such as a phrase or sentence. We have outlined in the introduction a framework for the study of memory in complex cognitive activities. We briefly review here the essential features of this framework as it is relevant to text comprehension, and then describe a model of text comprehension that incorporates this framework. Finally, we discuss some evidence supporting the two key assumptions we have made about the use of memory in text comprehension.

Comprehension involves a sequence of successive states. Most of the time we take as a comprehension state the thoughts arising from the processing of a sentence (or portion thereof, if the sentence is too long). However, a finer unit of analysis is sometimes indicated, and comprehension may be analyzed word by word. Thus, each state is characterized by its own textual input, and is a word or sentence, depending on the level of analysis. In the section above on within segment-processing, we were concerned with what we have called the vertical processing sequence and its memory requirements (see Figure 3). We are now looking at the horizontal sequence of states and ask the question "What information from previous states is available for processing in the current state?" We make the usual assumption that only information in working memory is available for processing, whereas long-term memory contents not in working memory must be retrieved and activated in working memory before they can be processed. We have argued before that two types of prior information are still available in working memory. One type is a small set of items which are still being attended to, that is, are maintained in a short-term memory buffer. The other type is a portion of the processing traces resulting from comprehension of the prior text. This portion is directly accessible from the current input and the prior attended information via the retrieval structures formed during comprehension.

The short-term memory buffer is independent of content and is a structural feature of memory with limited capacity. Alternatively, we could be speaking about maintaining information in the focus of attention, or about information that is directly accessible via temporal retrieval cues (recency cues). Not everything we have read can remain in the focus of attention[2]; temporal retrieval cues will be sufficient for the most recent text - a small fraction of what we have read.

A portion of long-term memory is directly available during the comprehension process - the extended or long-term working memory, as we have called it. This part of long-term memory is accessible via retrieval structures that were formed during the comprehension of the prior text. These retrieval structures are linked to retrieval cues in the input or in the prior text that is still in the focus of attention. The formation of effective retrieval structures depends on the comprehension skills, domain knowledge of the reader, or both. If successful, these retrieval structures turn a part of long-term memory into working memory space.

Thus, unskilled readers reading a text in an unfamiliar domain would have to depend primarily on the short-term memory buffer for the construction of a coherent textbase. Skilled readers in a familiar domain, on the other hand, ought to be able to rely mostly on using their long-term memory as an extension of working memory. A direct test of this hypothesis is not yet available. The discussion of Kintsch's (1988) construction

---

[2] It is of course not the case that people always attend to everything they are currently reading either. However, merely for the purpose of simplification, we make here the usual "diligent readers" assumption that readers do everything they are supposed to do.

integration model illustrated how a model of comprehension can incorporate such assumptions about memory.

The Short-Term Buffer in Comprehension.

In the Kintsch and van Dijk (1978) model and its successors, a limited-capacity, short-term memory buffer plays a crucial role in comprehension. By maintaining information in working memory, the short-term memory buffer makes it possible to construct a coherent textbase. A small number of propositions are selected at the end of each processing cycle and carried over in a buffer to be re-processed with the input propositions from the next processing cycle. The particular strategy for selecting these propositions suggested by Kintsch and van Dijk (1978) was a statistical approximation, based on predicted importance and recency. Fletcher and Bloom (1988), however, have done empirical work to investigate the selection strategies readers actually use. For certain texts, readers prefer to retain propositions that are likely to play a causal role. In the current version of the model (Kintsch, 1988), it is no longer necessary to specify selection strategies because propositions are selected on the basis of their activation values (which are determined by multiple sources, including causal relations, as in Kintsch, in press b).

According to the model, propositions that are retained in the buffer over one or more processing cycles accrue more strength in memory and enter into more relations with other text elements. Hence they will be recalled better. Researchers starting with Kintsch (1974) and Meyer (1975), have often observed the propositions of a text can be arranged hierarchically (based on linguistic, rhetorical, or other considerations, such as argument overlap) in such a way that the more superordinate propositions will be recalled better. This "levels" effect in recall is predicted quite well by the buffer model (Miller & Kintsch, 1980).

The size of the buffer in various applications of the Kintsch and van Dijk model has been estimated as between 1 and 4 propositions (Kintsch & van Dijk, 1978; Miller & Kintsch, 1980; Spilich, Vesonder, Chiesi, & Voss, 1979). This estimate agrees well with other estimates of short-term memory capacity. If all resources can be used for storage, as in a memory span test, about seven to nine chunks can be retained (Miller, 1956). If, on the other hand, subjects learn a word list for later free recall, a task for which most resources are devoted to encoding of information in long-term memory rather than short-term storage, only about two items are reproduced from short-term memory (Glanzer & Razel, 1974).

The reduced ability of older readers to reproduce a text has been attributed to a reduction in the size of the buffer available to these people during reading. Spilich (1983) has compared healthy older people more than 70 years old with college students. He found that the larger buffer capacity of college students (4 propositions) compared to that of older people (1 proposition) could account for differences in their recall. Interestingly, the model could not account at all for the pattern of recall obtained from a third group of subjects - older, senile persons. Healthy older people appear to use the same reading strategies as college students do, but are less efficient, whereas the poor memory performance of senile people appears to reflect an inability to develop an organized representation of the text base in memory.

Retrieval: Reinstatements and Inferences

We have argued that a limited amount of text is held in the focus of attention, that is, in working memory. The text held in working memory includes the most recently read text and some important information from the previous text carried over in a short-term memory buffer. When for some reason more information about the previous text is needed, for example to bridge a gap in the coherence between the previous text and the current text segment, it has to be retrieved from long-term memory. The needed information may be either a part of the already processed text (a reinstatement in working memory of the text), or general knowledge required to bridge some gap in the text (an inference).

Fletcher (1981) has investigated experimentally the availability of text elements in working memory and text elements that had to be retrieved from long-term memory. He showed that propositions predicted by the Kintsch and van Dijk (1978) model to be available in working memory are recognized more accurately and more quickly than propositions that need to be retrieved from long-term memory. In his experiment, propositions from the current processing cycle were always supposed to be available in working memory, whereas propositions from prior cycles always had to be retrieved from long-term memory. The crucial data came from the next to the last processing cycle, in which the model assigned some propositions to working memory and some to long-term memory. As Figure 9 shows, propositions from the next to last cycle assumed to be still available in working memory behaved more like propositions from the current (last) cycle, whereas propositions assumed to be in long-term memory behaved more like the propositions from prior cycles.

Insert Figure 9 about here

Anderson (1990) estimated the time required for a single retrieval operation at about 400 ms. From Figure 9 we obtain an estimate of around 300 ms. This is a quite reasonable value considering that the predictions of the model were undoubtedly not entirely correct, so that the results shown in Figure 9 can be regarded as approximate at best. Thus, the time it took to reinstate textual information in the Fletcher (1981) study is consistent with the idea that a retrieval from long-term episodic text memory is involved.

When readers make bridging inferences, they must retrieve information from general long-term memory, not from the episodic text memory. Kintsch and Keenan (1973) gave subjects sentence pairs like the following to read: "A burning cigarette was carelessly discarded. The fire destroyed many acres of virgin forest," or sentence pairs in which the first sentence contained the explicit information that the "cigarette started a fire." Subjects then verified test sentences like "A discarded cigarette started a fire." Reaction times were 400 ms (in another experiment with longer texts, 500 ms) faster when subjects had read the explicit statements than when they had to make the bridging inference on their own. On the other hand, when the test question was delayed for 20 min, there was no difference between conditions, presumably because long-term memory retrieval was involved in both cases.

The lack of a difference between explicit and implicit texts can be explained in two ways. One explanation is that subjects had made a bridging inference during reading of the implicit texts. The other explanation, and the one we favor here, is that the retrieval of the episodic text structure also retrieved the associated knowledge about the general world, so that the information that burning cigarettes may cause forest fires was available in working memory, whether stated explicitly in the text or not.

The Role of Domain Knowledge in Comprehension.

Everyday memory is so good because it is memory for meaningful materials, that is, materials for which people have developed efficient encoding strategies. Readers can reproduce texts very well if they have rich domain knowledge (Afflerbach, 1990; Bransford & Johnson, 1972; Dooling & Lachman, 1971; Recht & Leslie, 1988; Schneider, Körkel, & Weinert, 1989; Spilich, Vesonder, Chiesi, & Voss, 1979). Readers practice comprehension strategies all their lives and apply them rapidly, effortlessly, and unconsciously. When readers comprehend a text, they tie it to such knowledge structures as schemata, frames, or scripts (van Dijk & Kintsch, 1983). Such integration of new information and old knowledge structures is characteristic of expert memory. Memory structures like frames organize propositions in the text and also serve as effective retrieval structures (Kintsch, 1982).

Without adequate background knowledge, readers neither understand nor remember a text. For that reason stories are easy to comprehend because stories deal with people, their plans and actions - even children are already expert in that domain. On the other hand, as Bransford and Johnson (1972) showed, without knowledge, individuals cannot reproduce texts well at all. Bransford and Johnson wrote brief essays on familiar topics such as laundry, but in such a way that the topic could not be identified from the text itself. On the average, subjects reproduced a mere 3.6 words out of 14. In other words, they were no better off than if they had read a random word list. But if a title identified the topic for each paragraph, subjects' performance more than doubled to an average of 8 words. In the latter condition subjects could activate thematic knowledge and integrate it with the information provided by the text, thereby obtaining an effective retrieval structure.

Without domain knowledge and the skills to use that knowledge, a reader is described by the 1978 version of the Kintsch and van Dijk model. If the short-term memory buffer is insufficient to establish a coherent text representation, the reader must retrieve previous portions of the text or whatever bits and pieces of knowledge are available. Each retrieval requires a separate operation in the absence of an effective retrieval structure. This process can consume both time and resources and is a major source of reading difficulty (Britton & Gulgoz, in press; Kintsch & Vipond, 1979; Miller & Kintsch, 1980).

On the other hand, with domain knowledge and the skills to use it, readers are able to form retrieval structures so that whenever they need previous information from the text or relevant general knowledge, a single retrieval operation brings the appropriate portions of their long-term memory into working memory. Hence a major stumbling block for comprehension in an unfamiliar domain becomes trivial in a familiar one.

Consider the contrast between the following two sentence pairs: "John's car broke down. The motor just quit" versus "An abnormally low amount of hydrocele was found. The spermatic cord appeared quite dry." A bridging inference is required in both cases, but the inference is trivial in the first case: The knowledge that cars have motors and that the motor quitting is sufficient cause for a car to break down is readily available; "car" and "motor" are effective retrieval cues that bring the relevant parts of long-term memory into working memory, so that a coherence problem does not even arise. But "hydrocele" and "spermatic cord" retrieve nothing for most readers.

Either the sentence pair remains incoherent, or the reader engages in deliberate, conscious inference processes, hypothesizing, for instance, that because the spermatic cord was dry, it might be the place where hydrocele was low.

We have thus found evidence for a role in comprehension both for a content- independent, capacity-limited, structural, short-term memory buffer and for an expanded long-term working memory. Unfortunately, empirical studies clearly delimiting these two factors and examining their interaction have not yet been conducted on the topic of comprehension. There exist some suggestive results, however. For instance, Gernsbacher, Varner, & Faust (1990) found that skilled readers performed significantly better on a question-answering task that referred to the early part of a story but not when the question referred to the last part. Presumably, the skilled readers had an advantage because they had formed better retrieval structures to support retrieval from long-term memory, but that advantage played a lesser role when subjects could still rely on their short-term memory. This interpretation is consistent with Gernsbacher's (1980) observation that the skill advantage was present only when subjects were given normal stories (picture series), and not when they were given scrambled materials, which did not enable them to form adequate retrieval structures.

Alternative Accounts of the Role of Memory in Comprehension.

There are two alternative accounts to the notions about memory in text comprehension that we have sketched in this article. The first one argues that the surface form of just-preceding sentences is kept available in short-term memory to allow successful integration of the text. The second account proposes that the working memory capacity for maintaining relevant language information is relatively large and that individual differences in this capacity constitute a powerful predictor of successful text comprehension. We examine first the evidence for surface memory of previous sentences in short-term memory.

Short-term Storage of Linguistic Surface Form during Comprehension

We have distinguished a sequence of buffers for the temporary storage of intermediate results from the memorial processes that bridge the gap between successive states of thought: the vertical and horizontal dimensions in Figure 3. Thus, what is carried over from one state of thought to the next is fully interpreted structures - end products of comprehension rather than intermediate processing data. A popular alternative view, most clearly formulated by Glanzer and his colleagues (Glanzer, Fischer, & Dorfman, 1984;

Glanzer & Nolan, 1986), holds that what is carried over in short-term memory is the linguistic surface form of the sentence.

Investigators tried to resolve this issue by examining how much of the preceding sentences subjects can recall verbatim. If subjects are interrupted during reading at randomly selected places and asked to reproduce what they have just read verbatim, they reproduce the last two sentences almost perfectly, but not earlier sentences (Jarvella, 1971; Perfetti & Lesgold, 1977). This replicable finding seems to prove that readers maintain the surface form of about two sentences (or phrases). However, these findings merely show that readers in these experiments are able to reproduce most of the surface form of two sentences. Since the subjects in these experiments know that they will be tested for short-term retention, it is very likely that they use special chunking and rehearsal strategies, so that two sentences is almost surely an overestimation of the contents of short-term memory during normal reading, just as the immediate memory span is an overestimation of the short-term capacity during list learning (Glanzer & Razel, 1974) or reading (Daneman & Carpenter, 1980). In fact, when subjects are informed in advance of the way they will later be tested, their speed and pattern of reading are dramatically influenced by the particular test they anticipate (for example, Aaronson & Ferres, 1984; Kieras, 1984).

In our view, surface information remains active during the processing of a sentence. Afterwards this information is not directly available, although much of it can be reconstructed from the semantic encoding in long-term memory. Potter and Lombardi (1990) found the kind of synonym intrusion in sentence recall that would be expected from such a view of short-term memory. Consistent with this interpretation, Ericsson and Karat (Ericsson & Chase, 1982) found that errors in verbatim recall in a task measuring memory span for words in meaningful sentences were virtually always related to surface features irrelevant to the propositional content of the sentence.

Glanzer and his colleagues (Glanzer, Fischer, & Dorfman, 1984; Fischer & Glanzer, 1986; Glanzer & Nolan, 1986) performed an important series of studies to explore the role of short-term memory during reading. Their basic procedure was to compare continuous reading of a text to conditions in which the reading of consecutive sentences was interrupted at a given point in the text by some other task. This could be another reading task (reading some unrelated sentence) or a totally different task such as arithmetic. The most notable feature of their results was that this procedure, which appears extremely disruptive, had no effect on reading comprehension as measured by the subjects' ability to answer questions about the text. It did, however, significantly slow down reading times. In one of their experiments (Glanzer, Fischer, & Dorfman, 1984, Exp.1), subjects required 3,896 ms to read each sentence the text when the sentences were presented continuously, but needed 4,210 ms when an unrelated independent sentence had been interspersed between two consecutive sentences in the text. Interestingly, this was the case only when the sentences formed a coherent text. When the sentences were independent, reading times were much longer overall due to more words per sentence, but reading times were no different for consecutive indpendent sentences (6,6,90 ms) compared to independent sentences interspersed in the text (6,6,95 ms). The increase in reading time due to the interruption of an interspersed sentence (314 ms for Exp. 1, 364 ms across experiments according to Fischer and Nolan [1986]) is consistent with estimates reported in the introduction for retrieval from long-term working memory when the relevant retrieval cues are maintained in short-term memory. In another experiment (Fischer & Nolan, 1986, Exp.4) the reading of consecutive sentences in the text was interrupted by 30 s of arithmentic. Reading times for sentences directly following the interruption with arithmetic were aroung 1,700 ms longer than with the continuous reading condition. This estimate of retrieval from long-term memory is consistent with other estimates reported in the introduction for cases when relevant retrieval cues in short-term memory have been lost.

A critical result for Glanzer and his colleagues was that the they could eliminate the increase in reading times for the sentence following the resumed reading of the text by having the subjects first re-read the sentence they had read before being interrupted, then read the remaining sentences in the text. Glanzer and Nolan (1986) inferred from their results that the surface form of the last couple of sentences needs to be stored in short-term memory for normal comprehension of a text.

Jarvella (1971) and Glanzer's results can also be interpreted as indications that short-term memory maintains surface structure to allow the comprehension mechanism enough processing time and a chance to backtrack when necessary. Thus, the raw material of comprehension would remain available for reprocessing a little longer. There is ample reason to doubt this conclusion, however. What is carried over in short-term memory are end results, not the raw data or intermediate computations. Short-term memory contains fully identified elements along with associative connections, not the raw data from which these were derived. Cognitive operations in working memory are performed as soon as they can be; and it is the products of these operations that are carried over from one processing cycle to the next. In terms of our metaphor, what happens along the horizontal axis of memory is dependent upon the products of the vertical layers of processing.

Just and Carpenter (1980, Carpenter and Just, 1981) have argued for the immediacy of processing during reading on the basis of an analysis of eye movement during reading. Consider, for instance, the data reported by Carpenter and Daneman (1981). Using garden path sentences, they showed that the duration of gaze for ambiguous words is a function of the strength of contextual priming as well as of word frequency. In other words, readers encode, retrieve, and integrate an interpretation of an ambiguous word while fixating that word. Similarly, the word that disambiguated their garden path sentences was fixated longer when it was unexpected, and if found inconsistent, often gave rise to various attempts at error recovery. Thus, words are interpreted and semantically fixed immediately (or rather, as we have argued before, within 350 ms; see Figure 5).

Strong evidence for the immediacy hypothesis also comes from several studies on pronoun identification in discourse. Pronouns with ambiguous referents are typically fixated longer than pronouns with unambiguous referents (e.g., Vonk, 1985). Readers frequently look back when they encounter an ambiguous pronoun, usually to one or the other possible referents (e.g., Carpenter & Just, 1977). Like homophones for which readers access both possible meanings, but immediately (within 350 ms) select the appropriate meaning and deactivate the inappropriate one, (e.g., Frederiksen, 1981), readers also access all referent candidates for syntactically ambiguous pronouns but quickly decide on the appropriate meaning on the basis of semantic and discourse considerations. Finally, research on comprehension during listening (Cole & Jakimik, 1980) reveals that subjects detect mispronunciations at the first logically possible time, strong evidence for the immediacy hypothesis.

Of course, not all processing can be done immediately. For instance, the verbal instruction given to a patient, "Touch the green square with the red triangle" probably takes longer to understand than to hear. Hence some phonological buffering is necessary to be able to understand it (Baddeley & Wilson, 1988). Or, more typically, ambiguous words cannot be disambiguated until the end of a long sentence, so that either parallel constructions or the sound pattern itself must be carried along in memory for some time. McDonald and Carpenter (1981) have observed the eye movements of readers translating a text into another language. Translators try to translate word by word, as long as the text allows, but when problems arise (e.g., the presence of idiomatic, nonliteral expressions), they first chunk the text in English into a meaningful unit, and then access the second language to translate it. Similarly, readers must have available a sufficient chunk of a text for pronoun disambiguation. Indeed, Mathews and Chodorow (1988) have shown that when searching for the antecedent of a pronoun, readers search the text available in working memory in a top-down, left-right, breadth-first fashion. At the end of each sentence, readers typically give themselves a little wrap-up time to catch up with whatever processing could not be performed during the reading of the sentence, such as determining the sentence topic (Aaronson & Scarborough, 1976).

This view of text processing necessitates a slight reinterpretation of Glanzer's results. Glanzer showed that if short-term memory is disturbed after each sentence of a coherent text is read, understanding does not suffer but reading times increase. During normal reading, the integration of the current sentence with the previous text is facilitated because the most activated elements of the previous sentence are still in the focus of attention and hence can serve as a bridge between the long-term memory trace of the text and the currently processed sentence. If this focus of attention is disrupted, as in Glanzer's interruption procedure, the reader must retrieve and activate the long-term memory representation corresponding to the previously processed text to successfully integrate the information in the new sentence. Typically, this retrieval operation requires about 400 ms, a number that agrees well with other estimates for a single long-term memory retrieval. Once the retrieval of the organized long-term memory representation of the text is completed, subsequent retrieval of more specific information (cf. anaphoric reference) can be made by relying on associational connections with the activated elements in attention (short-term memory). The interruption procedure, therefore, should do little more than slow down the reading process a bit. What is disrupted is neither a perceptual trace nor purely a surface representation of the text, but a fully analyzed, fully interpreted representation of the previous text.

The claim that the products of comprehension are maintained in short-term memory (and stored in long-term memory) should not be confused with the claim that memory for text is necessarily propositional and that linguistic surface structures are not retained. Short-term memory contains products, but what these products are depends on many factors. Often they are semantic representations, but they may be spatial images (as in the situation models studied in Perrig & Kintsch, 1985) or some other, nonpropositional mental structures. Sometimes the products are indeed surface structures, as in the recitation of a children's counting rhyme (Kintsch, in press), or in remembering the particular nuances of a conversation.

A well-known case in which memory for text was almost purely situational is found in the experiments reported by Bransford & Franks (1971). Subjects read sentences which were quite similar in their surface structure ("The ants were in the kitchen. The ants ate the jelly.") but allowed them to construct a simple mental model of the situation. In this case, subjects had excellent memory at the level of the situation model, but no surface memory, presumably because of the strong interference between sentences at that level of representation. In contrast, in many laboratory experiments using independent sentences, the representations formed by the subjects appear to be primarily propositional. The exact surface form is too irrelevant to the task at hand, as well as too difficult to be remembered.

It is also clear that surface features may persist in long-term memory. At one point, the results of some early studies were misinterpreted, researchers concluding that the surface features of the language were nothing but the chaff to be discarded as soon as the precious meaning was extracted. Although it is generally true that meaning is retained better than surface memory (Sachs, 1967; Bransford & Franks, 1971; Hanson & Bellugi, 1982, for sign language), long-term retention of surface form is by no means rare (Kintsch & Bates, 1977; Masson, 1984; Hjelmquist, 1984). Indeed, surface form is retained best when the way something is expressed is pragmatically significant. It matters a great deal whether a partner in a discourse has said something politely or aggressively, and in these situations the wording is quite well remembered (Bates, Kintsch, & Fletcher, 1980; Keenan, MacWhinney, & Mayhew, 1977). However, in laboratory studies of memory for sentences, outside a social context, memory is typically propositional and surface features are indeed reconstructed (Potter & Lombardi, 1990).

Capacity of the Transient Portion (Attention or Short-Term Memory) of Working Memory

Our theoretical framework distinguishes between information about the previous text stored in retrievable form in long-term memory and the information kept active in attention or short-term memory. An important function of the information maintained active is to serve as retrieval cues for efficient retrieval of the information stored in the long-term memory portion of working memory. Because most other theories of working memory consider only the temporary storage of information in attention and short-term memory, we will re-examine some of the basic results associated with these alternative theories to determine the extent to which storage in the long-term memory portion of working memory could be involved. We first discuss the impact on reading of concurrent memory tasks demanding additional temporary storage of information. Then we discuss efforts to measure individual differences in working memory for language information.

Concurrent cognitive processes compete for the resources of working memory. Baddeley and his colleagues have shown that this competition may decrease performance for paired-associate learning, free recall, and text comprehension (Baddeley, 1986; Baddeley & Hitch, 1974). When subjects are asked to hold six digits in memory while reading a sentence and later to reproduce those digits, a decrement in performance of about 10% is obtained compared to the performance of subjects without memory load. In another experiment, subjects had to memorize three or six visually presented digits while listening to a prose text. With only three digits to learn, subjects answered questions about the text almost as well as a control group (4% less). With six digits, however, a significant performance decrement of 18% was obtained.

In other experiments Baddeley (1986) found performance decrements for comprehension as well as for memory. In one of these studies, subjects had to verify visually presented sentences (as in semantic memory experiments) while remembering from zero to eight spoken digits. Figure 10 shows that the frequency of error increased when subjects had to remember six or more digits, and that reaction times generally increased with memory load.

Insert Figure 10 about here

Results like these show that in dual task situations, text comprehension is relatively unaffected by low and intermediate loads on memory and only really impaired when the resource demands of the secondary task are maximal: even without anything else to do, people can manage to remember six or eight digits at most. As long as they have some free resources, however, they perform remarkably well on text comprehension tasks. This finding should be expected because text comprehension is a highly overlearned, expert skill. In comprehending text, people function quite well with an expanded, long-term working memory and are not forced to rely on resource-consuming, short-term memory maintenance. That, at least, would be the case for simple texts. For difficult, unfamiliar texts the short-term buffer should be expected to play a more significant role.

Other investigators have inferred from these and similar results that successful text comprehension might depend on the availability of sufficient working memory capacity. The capacity of working memory may be larger for good readers than for poor readers. Numerous studies of short-term memory seem to contradict this claim. Memory span for children older than 5 years is the same as for adults (Case, 1978; Chi, 1976; Dempster, 1981; Huttenlocher & Burke, 1976); there is no difference in memory span for good and poor readers (Farnham-Diggory & Gregg, 1975; Rizzo, 1939). The memory span, however, is purely a test of storage capacities and does not indicate the capacity of working memory available during reading.

Working memory, on the other hand, has a dual function, processing as well as maintenance. If poor readers need to devote most of the capacity of working memory to decoding processes, the capacity for maintaining information becomes limited (Case, 1978; Daneman & Carpenter, 1980, 1983). Daneman and Carpenter (1980) designed a task to measure the capacity of working memory during reading. They present subjects with a series of unrelated sentences, which they need to comprehend to answer subsequent test questions. At the end of the presentation, subjects are asked to recall as many of the last words of the sentences as possible. The number of words a subject correctly recalls is referred to as that subject's reading span. College students manage to recall the last word from five to six sentences; good readers have a significantly higher reading span than poor readers. Reading span correlates with comprehension tests (r=.5 to .6) and with the ability to answer content questions (r=.7 to .9). Individual differences among readers are related to reading span: In Daneman & Carpenter (1983) 27% of the variance is accounted for by the reading span measure compared with only 10% by differences in reaction times.

Insert Figure 11 about here

Figure 11 shows, for readers with reading spans from 2 to 5, the likelihood that they can correctly identify an ambiguous pronoun as a function of the number of sentences intervening between the pronoun and its referent. As the number of intervening sentences increases, working memory capacity becomes more and more crucial. Readers with a low span fail completely, whereas readers with a high span have no trouble at all. Daneman and Carpenter (1980, 1983) proposed that the capacity of working memory is constant, but that good readers have more room for storing additional information because less of their working memory is taken up by the reading task due to their higher skill level. According to this account, the last words of sentences are maintained in active form in working memory.

In our framework the most natural interpretation of the reading span test is that it measures subjects' ability to store information about sentences in the long-term memory portion of working memory. Because the subjects can correctly answer questions about all the sentences, we know that they must have formed a representation in long-term memory. By storing additional information about the sentence, they can later retrieve the final word of the sentence. In addition, subjects must be able to encode associative connections to allow cued recall of the sentences and their last words.

Our account leads to several empirically testable predictions. First, further testing of subjects' memory of sentences in the reading span test should reveal long-term memory representation of these sentences. Masson and Miller (1983) found that cued recall of other words of the sentences were as good as predictors of reading comprehension as the final words were. Baddeley (1986) found similar correlations between comprehension ability and a modified reading span test in which subjects were told only after the end of presentation what type of information they had to recal. In this version of the reading span task, the subjects could not anticipate what information would be requested and therefore had to maintain a lot of information for successful performance. The storage of large amounts of information is consistent only with storage in long-term memory.

Second, the ability to remember sentences should be related to verbal ability and the ability to comprehend text. Masson and Miller (1983) found that delayed testing of recognition memory for explicit and inferred statements from a paragraph were both as highly related to text comprehension performance as to the reading span scores. Ericsson and Karat (Ericsson & Chase, 1982) found that memory span for words in sentences was highly correlated to a test of verbal ability.

Third, reading span is also correlated with the ability to make inferences, although if one partials out subjects' performance on questions asking for information stated explicitly in the text, this relation is no longer significant (Masson & Miller, 1983; Dixon, LeFevre, & Twilley, 1988). Singer, Andrusiak, Reisdorf, and Black (1989) have shown, however, that this correlation depends on the type of inference. Subjects are likely to make bridging inferences as an integral part of constructing a textbase because these inferences are necessary to make the textbase coherent. The correlation of bridging inferences with reading span remains significant even after explicit memory is partialled out, at least when the memory load is substantial (premises are separated by three intervening sentences). On the other hand, subjects are likely to make deductive inferences in response to the question rather than on-line. Deductive inferences are independent of working memory capacity once explicit memory is accounted for.

The high correlations between text comprehension and, on the one hand, measures of memory performance and, on the other hand, tests of knowledge of language (vocabulary and grammar) is consistent with the assertion that text comprehension is an acquired skill. An important aspect of this acquired skill is storage of an integrated representation of the previous text in long-term memory. The storage itself must be rapid and accurate, and it must allow efficient retrieval of this information whenever needed. We suggest that Daneman and Carpenter's reading span measures this ability to store and later retrieve information about preceding sentences from long-term memory. Thus, what we are dealing with in the studies we have reviewed is not maintenance of temporary information in working memory, but skilled readers' ability to access long-term memory from retrieval cues held in the active portion of working memory.

Some of the best evidence for our notion of long-term working memory and against inherent individual differences in temporary working memory capacity comes from research that systematically varies verbal ability and relevant domain knowledge. Recht and Leslie (1980) selected four groups of subjects on the basis of their reading ability (high and low) and their knowledge about baseball (high and low). Schneider, Körkel, and Weinert (1989) similarly selected four groups based on aptitude (high and low IQ) and knowledge about soccer (high and low). Both studies found that memory and comprehension of texts describing events in soccer or baseball was influenced only by the amount of knowledge (high-low). There was no evidence for a main effect of or an interaction with IQ or reading ability. Hence, students with low reading ability and expert knowledge clearly outperformed students with high reading ability and little knowledge. The surprising absence of any effect of reading ability led Recht and Leslie (1988) to propose that the superior reading ability observed for standard texts, where reading ability is related to comprehension and memory, is mediated at least in part by differences in amount of knowledge relevant to such texts.

Conclusions: Memory for Text

The vertical dimension of text processing requires a series of buffers for temporary retention of the intermediate computations involved in comprehension-- beginning with the various stages of the perceptual analysis and continuing to the higher cognitive process--such as semantic interpretation and the construction of a situation model. At the lower levels of this processing chain these buffers are specialized, dedicated neural structures. At the upper levels, any demands for retention must be satisfied within the general constraints of working memory.

We have characterized thinking in general as a succession of different thought states, which in the domain of text comprehension can be identified with the sequence of processing cycles in the Kintsch and van Dijk (1978) model. We are now dealing only with the top of the vertical processing hierarchy involved in comprehension. Even at that level, however, a great many intermediate computations are performed and lead to results that are later discarded. In Kintsch's (1988) construction-integration model, for instance, a great deal of activated knowledge turns out to be contextually irrelevant and is quickly disposed of. Multiple interpretations of ambiguous input are made, and associated knowledge is retrieved indiscriminately, but much of this turns out to be irrelevant or even contradictory and is therefore deactivated in the integration phase of each comprehension cycle. It is the final, integrated representation, comprising in various degrees surface features, propositional structures, and a situation model, that is stored in long-term memory and maintained in activated form in short-term memory between processing cycles. The long-term memory representation is complete--presumably, everything that survived the integration process is stored, although it may be forgotten later. The short-term memory, on the other hand is highly selective--only the momentarily most activated elements of the text representation are preserved in the focus of attention from one processing cycle to another.

A review of the evidence presented in this article shows that the way memory is used in text comprehension is described quite well in most respects by conventional memory theories derived from laboratory studies. For instance, although we have not specified actual retrieval mechanisms here, current models of memory retrieval (e. g. Raaijmakers & Shiffrin, 1981) offer likely candidates. Nor does our notion of short-term memory deviate very far from classical conceptions. What is different, however, is the concept of the long-term working memory: a rich store of information connected via retrieval structures to cues in working memory, a store that, while not active itself, is directly retrievable and hence functions as an extension of working memory.

Insert Figure 12 about here

Figure 12, an elaboration of Figure 4, further clarifies the notion of an expanded long-term working memory. The sequence of buffers involved in the transformation of the text input is shown at the bottom of the figure. The diamond labeled CONSTRUCTION stands for a complex system of production rules that derive propositions from the linguistic input, form situation models, and add all kinds of associated knowledge. The INTEGRATION process rejects those elements created by these production processes that do not fit into a coherent whole. The result is the text representation for the current processing cycle, *n*. This representation may include propositional, situational, as well as surface components. Also active in working memory is a small portion of the text representation created on the previous cycle, shown by the darkly shaded area. This is the short-term memory buffer, in which a few of the most important text elements from previous cycles are maintained for further processing. Another portion of the memory trace from the previous processing cycles (the lightly shaded area) constitutes the expanded, long-term memory part of working memory. While no longer active (that is, in the focus of attention), it is directly retrievable via cues in the active part of working memory and hence forms, effectively, an extension of working memory. This extension also includes a portion of general long-term memory, that is, relevant knowledge, personal experiences, or both, associated with the text that is currently being processed. Other parts of long-term memory (the unshaded parts), both previous text and other information, are not directly accessible from working memory. Access to those parts requires problem solving activities to construct appropriate retrieval cues. For skilled readers reading a text in a familiar domain, the lightly shaded area would be large. For unskilled readers in an unfamiliar domain, this area would be small or negligible.

We now turn from text comprehension to problem solving and other cognitive activities in which the moment-to-moment course of processing is less constrained by an ever varying, continuous flow of inputs. We attempt to show that in these processes, direct access to information in long-term memory plays the same role as in text comprehension and depends upon the same factor, namely the availability of retrieval structures that characterize expert performance.

Memory in Problem Solving and Other Cognitive Activities

It is both easy and difficult to extend our theoretical framework of long-term working memory to empirical studies of problem solving and other types of complex cognitive activities. It is relatively easy in that many findings, as discussed in our introduction, can be adequately accounted for by the classical model, in which the cognitive processes are represented by a sequence of states of information in short-term memory. However, the major reason for the good fit of the classical model to the experimental results is that researchers have deliberately searched for laboratory tasks that minimize the role of long-term memory in successful completion of these tasks. In this section we expand the discussion to include more complex tasks in which retrieval from and storage in long-term memory plays an increasingly important role. We also discuss cognitive processes in experts, where evidence on the use of long-term working memory is available.

In our view, the sufficiency of the classical model of working memory in accounting for many results on problem solving and other complex activities is due to careful selection of laboratory tasks and stimuli to elicit these cognitive processes. Experimental psychologists have traditionally attempted to design laboratory tasks and situations that allow them to study a particular mental function in its purest form. Hence they have tried to minimize the role of memory in their studies of problem-solving, thinking, reasoning, and decision making. In earlier studies, investigators tended to select tasks domains that were unfamiliar to subjects and hoped thereby to eliminate the influence of specific, pre-existing knowledge on the cognitive processes under observation. This means that the information necessary to perform the tasks has to be presented to the subject. Investigators relieved the demands for memory storage in most studies by letting the information remain perceptually available during the task. Newell and Simon (1972) referred to perceptually available information as being accessible from external memory. Hence we argue that most unfamiliar tasks used in laboratory research have been so designed that their execution minimizes demands on memory load. We examine the memory demands for these tasks later in this paper.

In addition, we argue that several characteristics of these tasks make use of long-term working memory for storage difficult if not impossible. By definition the stimuli used in an unfamiliar task does not correspond to pre-existing knowledge, which facilitates encoding and storage in long-term memory. Furthermore, the stimuli are often combinations of independent features which would maximize interference for encoding into and retrieval from long-term memory of stimuli subjects have seen previously. Finally, when subjects are unfamiliar with a task, they do not know what aspects of the information are important. Their efforts to selectively store information with appropriate retrieval cues is therefore severely hampered. However, as subjects gain more experience with initially unfamiliar tasks, we can see that they rely increasingly on retrieval from long-term memory. We discuss these results along with other findings on cognitive processes in familiar task domains, where retrieval of knowledge from long-term memory is critical. We view expertise and expert performance as an extreme case of efficient retrieval of relevant knowledge, and we predict that in this case, the use of long-term working memory can be observed.

In presenting research on problem solving, reasoning, and decision making and their relation to memory, we first discuss laboratory studies in which the demands on memory were minimized. We contrast these findings with research on subjects with considerable experience and expertise in comparable tasks. In particular, we consider how necessary information is stored during the cognitive activity and whether relevant information can be readily retrieved from subjects' long-term memory. We start with the classic work on problem solving by Newell and Simon (1972) and with other problem-solving research in which subjects were not familiar with the research tasks or the task domain.

Problem Solving in Unfamiliar Task Domains

When subjects are solving problems in unfamiliar task domains and when they lack knowledge about specific solution procedures, we would expect a clear separation between their efforts to understand the problem situation and their subsequent generation of the solution. Therefore, we first consider the phase of problem-solving involving storage and comprehension of information about the problem situation.

Most laboratory studies of problems in unfamiliar domains have certain design characteristics that dramatically reduce subjects' memory load. First, the problem situation is presented in a perceptually available version that can often be directly manipulated. In Figure 13 some well-known examples of problems are shown with their respective representations.

Insert Figure 13 about here

Second, the experimenter's verbal instructions to the subject are often quite short and help to structure the situation by stating what is given and what needs to be attained. Examples of the critical parts of instructions are also shown in Figure 13. Third, the goal state is almost always very easy to remember, for example, "The tiles are ordered from lowest to highest number in left to right order," or "All the discs on the right-most peg," and the goal state corresponds closely to pre-existing knowledge. If the goal is relatively difficult to remember, the description of the goal state is kept perceptually available, as it is in the logic task in Figure 13. The same is true for legal operations and transformations. Finally, as an additional aid for remembering constraints on the generation of the solution, an experimenter constantly monitors the legality of operations and reminds the subject about any attempts to violate constraints of the problem situation. In computerized presentations of these problems, the computer programs monitor the legality of moves. By all these methods, requirements for comprehension and storage of information about the problem situation have been minimized for problems in unfamiliar task domains. As would be expected, the duration of the phase of understanding the problem situation is relatively brief for these problems.

When Hayes and Simon (1974) decided to study how subjects construct the representation of the towers of Hanoi problem, they did not use the physical analogue shown in Figure 13. Instead, they presented a verbal description of a problem isomorphic to the towers of Hanoi and monitored the subjects' comprehension processes by having the subjects think aloud about the description. When they presented the problem situation incrementally, Hayes and Simon (1974) found evidence that subjects constructed a representation of the problem that followed quite closely the structure of the information in the problem description. A computer simulation model designed to create a representation of the problem description also followed the structure of information in the text to find the representation of the problem. In subsequent research Hayes and Simon (1977; Simon & Hayes, 1976) studied understanding of problem solving in many isomorphs of the towers of Hanoi. Their most striking finding was that almost without exception the structure of the presented problem description directly determined the subjects' representation of the problem. This finding appears to hold when subjects solve problems in unfamiliar task domains.

Generation of Solutions

To problems like those in Figure 13 subjects must transform the current state, which is perceptually available, into the goal state, which is often stored in memory, by means of certain legal operations or moves. To directly reach or even approach the goal state, subjects must select the correct sequence of operations. From analyses of think-aloud protocols, Newell and Simon (1972) found that subjects tried to reach the goal state by eliminating perceived differences between the current state and the goal state one at a time. Having selected a difference, a subject would focus on finding operations (means) to eliminate that difference (ends), and when successful would turn to one of the remaining differences between the current state and the goal state. This means-ends analysis has been shown to be the major method for solving problems in unfamiliar task domains; empirical evidence for its use comes from studies of many different types of problems (Greeno & Simon, 1988).

The relation of means-ends analysis to demands on memory is particularly interesting. It is important to note that the differences identified by the subjects correspond to the attributes defining the goal state. In the towers of Hanoi a subject would report, "The largest disc should be on Peg C", in the 8-puzzle, "The tile with number 1 should be in its place", and for the logic problem, "The negation sign in front of the expression should be removed." Hence, the differences can be detected with a simple matching procedure comparing the current state and the goal state. By focusing on only a single attribute or difference at a time, the subject minimizes the amount of information in short-term memory and attention. Focus on a single difference also simplifies the selection of a single operation or sequence of operations leading to the elimination of the difference. This is particularly true in the logic task (Newell & Simon, 1972), for which the number of possible transformation rules is relatively large. The information necessary for guiding the process of generating the solution is quite limited and appears to be within the capacity limits of short-term memory (Newell & Simon, 1972). However, subjects update the current state for these types of problems on a perceptually available display, which serves as an external memory.

Newell and Simon (1972) discussed rather extensively the characteristics of storage in and access from external memory. They argued that there are clear parallels between external memory and long-term memory. Access from external memory requires that the eyes be fixed on the corresponding information, and unless the subject has stored in long-term memory the location of the corresponding information, the subject must determine its location through visual search. Newell and Simon (1972) noted that considerable evidence for visual search through the transformation rules in the logic task was available in the beginning of the test session; but with more experience and practice, the subjects could directly access the location of rules with certain characteristics. Through this access, subjects could retrieve lower-level details or check the representation in long-term memory. In the later parts of practice, subjects accessed at least the more frequently used transformation rules directly from long-term memory. In sum, Newell and Simon (1972) argued that in problem-solving tasks like the logic tasks, storage in long-term memory is time-consuming and the amount of information stored is quite restricted.

Additional evidence that problem solving processes are constrained by available memory capacity comes from the observation that subjects rarely plan out mentally what to do, and if they plan, the depth of their planning is limited (Atwood, Masson, & Polson, 1980; Ericsson, 1975; Newell & Simon, 1972). Newell and Simon (1972) observed one of the few exceptions to the limit on planning with a subject on the logic task. This subject planned out his solution using a very abstract representation of the problem and of the perceptually available transformation rules, thus minimizing the memory load. Another empirical result, that subjects show very limited transfer from their first successful solution to their next solution attempt on the identical problem, is also consistent with the view that long-term memory for the generated solutions is quite limited (Karat, 1982; Reed, George, & Banerji, 1974, Exp. 2).

Problem Solving in Familiar Task Domains

In a familiar task domain subjects have a large body of relevant knowledge and experience. The amount of relevant knowledge provides a rough measure of the degree of familiarity. An expert with a vast amount of knowledge defines one extreme; a subject in an unfamiliar task domain defines the other extreme. Between these two extremes are many intermediate degrees of familiarity. Familiarity is normally a function of the relevant experience a subject has had with tasks in the domain. However, many unfamiliar tasks or task domains used in laboratory research become familiar to subjects after only limited experience. Within a single test session subjects' representation of the task is often changed (Karat, 1982). Occasionally the representation changes radically (Ericsson, 1975), and subjects discover new and quite different solution strategies (Anzai & Simon, 1979). Subjects can master the tasks in familiar domains by acquiring a single general solution method that can then be successfully applied with a minimal amount of specific knowledge. In contrast, most task domains in real life require access to large bodies of very different types of knowledge, which people acquire over extended periods of time.

In the introduction we argued that a necessary pre-condition for being able to use long-term working memory is the direct accessibility of a large body of relevant knowledge to allow rapid storage in long-term memory. As the following sections show, storage of knowledge in long-term memory does not mean that this knowledge can be rapidly accessed when needed. Beginners or novices with limited experience frequently have difficulty accessing relevant knowledge even when they deliberately attempt to retrieve it. In direct contrast experts seem to be able to retrieve relevant knowledge automatically, in a fashion similar to unintentional retrieval of information (reminding) (cf. Norman & Bobrow, 1979; Schank, 1982). On a theoretical level, research must explain how the relevant information is selectively accessed from the vast amount of information in long-term memory.

Retrieval of information during problem solving is much more difficult than retrieval during text comprehension. A well-written text carefully guides the readers' cognitive processes, presents the information that should be integrated, and provides good retrieval cues for necessary retrieval of information from long-term memory. The situation is very different for the problem solvers, who, once they have read and understood the problem description receive neither information to guide their cognitive processes nor good retrieval cues for the retrieval of the auxiliary knowledge necessary for generating a solution. Auxiliary knowledge includes facts and formulas, previous solutions to the same or similar problems, acquired solution procedures, and solution methods.

In this section we first consider evidence for the difficulty novices have retrieving relevant information and solution methods. We discuss studies that compare the problem solving of novices and experts in the same task domains, with particular emphasis on the mechanics of retrieving knowledge and solution methods. We then review the evidence for the use of long-term working memory by experts.

Retrieval of Knowledge

In familiar task domains, one would expect relevant knowledge to be readily retrieved when needed. However, for subjects who are novices, this access appears to be difficult and often unsuccessful. According to models of retrieval from long-term memory, access of the relevant information requires attention to retrieval cues, which are associated with the relevant information. The first step towards successful retrieval is the realization that something relevant can be retrieved. In a classic study, Bloom and Broder (1950) found that many subjects simply gave up on problems when they could not directly retrieve an answer from memory. For example, if asked to rank order the emergence of life forms such as fish, flowering plants, amphibians, and mammals, hardly anyone would have the correct answer memorized. However, most adults know quite a lot about these biological categories and can retrieve and use this knowledge to produce the correct answer by reasoning. Bloom and Broder (1950) found that the poor problem solvers knew enough information and that, if taught to retrieve and use it to find the answer by reasoning, could dramatically improve their test scores.

Subsequent research has shown that training in the use of related knowledge to search for a correct answer especially benefits students of low ability (Raphael & Mckinney, 1983). More generally, investigators (Walker & Kintsch, 1985; Whitten & Leonard, 1981; Williams & Holland, 1981; Williams & Santos-Williams, 1980) have studied subjects recalling members of different categories, such as names of teachers, names of students in a high-school graduation class, car models, and brands of detergent. These studies show consistently that subjects retrieve only a few members of the class from the name of the class. They retrieve a much larger number of members by generating more specific retrieval cues. For example, subjects recalling names of cars reported thinking of "cars in my dormitory parking lot", and "cars I wrecked" (Walker & Kintsch, 1985). When subjects are asked to recall, for example, "a time when you went shopping and couldn't pay for the item you wanted," they can retrieve such episodic experiences only after active generation of more specific retrieval cues (Reiser, 1986). In contrast to these relatively slow and disconnected recall activities, other recall tasks lead to fluent retrieval of elements. Kintsch and Mannes (1985) asked subjects to recall what typically happened when they went to a restaurant for a meal and to respond in other script-based activities while thinking aloud. Recall in these tasks was remarkably fluent and all the subjects recalled the activities in the same order, that is, the order in which they normally occur. A detailed analysis of the subjects' protocols revealed distinct episodes and clear organization of the script-based activities. However, the fluent recall of script-based activities is due to the continued availability of temporal retrieval cues leading to efficient retrieval of the next item throughout recall (Kintsch & Mannes, 1985). Hence, the memory representation of script-based activities appears not to be different from that of other complex knowledge structures. In all these cases, subjects have relatively well-organized knowledge systems, which allow them to generate the more specific retrieval cues necessary for successful retrieval.

Subjects' difficulties in retrieving knowledge relevant to task performance is so well known that it is rarely documented. The unreliability of retrieval of relevant knowledge is nicely demonstrated by people's ability to detect their own errors. Allwood and Montgomery (1981, 1982) have studied subjects detecting errors in their solutions to problems in statistics. They found that subjects were able to find some errors spontaneously without a directed effort to verify their solutions (Allwood, 1982). Subjects could detect more errors if the experimenter focused their attention on certain problem-solving steps. The principal factor in retrieving the correct knowledge was a broader and more intensive retrieval effort than occurred during normal problem solving. It is important to note that these subjects were not skilled in statistics and that their knowledge of statistics was not well integrated.

In several problem-solving domains, investigators have shown that increased experience with problems in the domain enables subjects to integrate knowledge and procedures to alleviate the problem of retrieval. Beginning students of theorem proving in geometry determine whether a given theorem is applicable by examining the preconditions for one theorem at a time. With more experience, students consider what information is given in the problem and on that basis retrieve appropriate theorem directly from memory through a recognition process (Greeno, 1976). With practice, sequences of steps in a procedure, which initially required attention for individual retrieval and execution, are integrated into cognitive units through processes of proceduralization and composition (Neves & Anderson, 1981).

Retrievability of Previous Solutions

Several studies are relevant to novice problem solvers' retrieval of previous solutions in order to solve other related problems. In learning to solve problems in new task domains, subjects appear to spontaneously retrieve solutions to earlier problems (Anderson, 1987b). This retrieval could be mediated by several factors such as the similarity of the content or surface structure of the problem, the similarity of the solution structure, or subjects' intentional retrieval of solutions in the belief that sequentially presented problems have the same solution structure. Ross (1984) varied orthogonally the similarity between previously solved problems and the target problem and monitored the frequency with which subjects retrieved previous solutions. He found that similarity of content or surface structure guided the retrieval regardless of the similarity of solution structure. Hence, similarity of surface elements appears to guide spontaneous retrieval of previous solutions encountered in the same session.

Studies by Gick and Holyoak (Holyoak, 1985) showed that only around 20% of the subjects spontaneously used the solution described in a previously presented story to solve an analogous problem. When explicitly instructed to use that solution, an additional 45% of the subjects were able to use the analogy. Subsequent work by Holyoak and Koh (Holyoak, 1985) showed that spontaneous use of the analogy could be substantially increased to 59% by making the surface elements of the previous solution more similar to the target problem. These results strongly implicate the role of similarity of surface features for spontaneous retrieval of earlier solutions. Holyoak (1985) also investigated the importance of presenting the solution and the analogous problem in the same session. Such a presentation would provide some demand characteristics that might prompt subjects to try to relate the described solution and the problem. However, Holyoak and Koh (Holyoak, 1985) were able to show that for problems with similar surface elements, students would spontaneously access and use solutions described in their psychology texts.

In a similar vein, researchers have studied the role of general methods in problem solving. Most studies have tried to determine how frequently Polya's heuristic methods are used and to account for individual differences in mathematical problem solving in terms of access and use of these general methods (Kilpatrick, 1968; Webb, 1975). The results of these studies have been disappointing. They show no relation or very weak relations between use of heuristics and problem-solving ability when the overpowering effects of general knowledge of mathematics (assessed by standard tests) have been controlled for (see Ericsson & Simon,1984, for a brief review). Studies instructing and training students in the use of Polya's heuristics have found marked transfer of the cognitive processes when the content of the training problems closely corresponded to the transfer problems. However, no transfer to problems with different content has been found (Lucas, 1972; Schoenfeld, 1979).

The results from studies in which analogous solutions and general problem-solving methods were retrieved consistently show the difficulty of retrieving such mental structures in the absence of common surface elements or detailed features. Considering the vast number of structures in long-term memory that could possibly pertain to the solution of the problem, a reasonable degree of match must be necessary for successful retrieval. However, for problems with additional constraining contextual cues that are known to the subjects, retrieval with weakly associated cues is quite possible.

Expert Problem Solving. When subjects have more extensive experience of problems in a given domain, there is evidence for a more efficient mode of problem solving that is commonly referred to as expert problem solving. Expert problem solving is generally characterized by an initial encoding of the presented problem in such a manner that solution methods can be directly retrieved from long-term memory. Hence, encoding of new problems is efficiently coordinated with the representation of solution methods and solutions of previously experienced problems stored in long-term memory.

Algebra word problem solving. The pioneering work on solving algebra word problems (Hinsley, Hayes, & Simon, 1978; Paige & Simon, 1966) showed that experienced problem solvers could use a range of solution methods. Subjects were able to solve meaningless problems using step-by-step translations into algebra. For meaningful problems, Paige and Simon (1966) were able to show that some subjects processed the problem text semantically and were thus able to detect logical inconsistencies in the problem situation. Hinsley, Hayes, and Simon (1978) demonstrated that subjects easily solved many problems by recognizing the type of presented problem and that subjects could reliably categorize problems into types without solving them. On the average, subjects required only the first sentence or first couple of sentences of the problem text to recognize the problem type. From these first sentences describing the context of the problem, subjects could often accurately predict subsequent information from the problem text as well as the appropriate solution procedures. To classify the relation between the problem context and the structure of the solutions, Hinsley, Hayes, and Simon (1978) constructed new problems, in which a familiar solution structure was embedded in an unusual problem context. Subjects occasionally remarked on the structural similarity of these unfamiliar problems although they did not seem to capitalize on their preexisting knowledge for solving problems with the same structure

Subsequent studies have tried to assess novices' and experts' representation of algebra word problems by having these subjects categorize problems in terms of similarity. Novices tended to categorize problems by their content (Silver, 1981); and when content was similar across problems, novices were unable to identify problems with the same solution structure (Berger & Wilde, 1987). Experts and problem solvers of high ability categorized problems according to their solution structure regardless of the cover story (Berger & Wilde, 1987; Silver, 1981).

The ability of experts to use the structure of the solution as a cue for categorization implies rapid generation of, or access to, procedures for attaining the solutions to problems. To gain better insight into subjects' comprehension of problems, Robinson and Hayes (1978) had subjects read algebra word problems statement by statement and indicate what information would be relevant to the solution. They found that good problem solvers were quite able to identify relevant information and that knowing in advance the question asked in the problem improved their ability to reject irrelevant information.

Physics problem solving. Novices' solving of physics problems can be described as means-ends analysis: Subjects start with the requested answer and search for a sequence of formulas that, with entities given in the problem, can be used to calculate unknowns and thereby derive the answer (Larkin, McDermot, Simon, & Simon, 1980; Simon & Simon, 1978). Novices' method of working backwards from the desired answer to generate a solution plan appears to be completely missing in experts. After having read the physics problem, experts proceed to generate the intermediate steps of the solution directly, that is, by working forwards.

Novices' and experts' representation of physics problems has been most systematically studied by Chi, Glaser, and Rees (1981). Experts efficiently categorized problems on the basis of the underlying physical principles. This behavior implies rapid access to a solution method for the problems. Novices, on the other hand, grouped problems predominantly according to shared situational elements, like problems involving pulleys and inclined planes. Experts' richer and more integrated representation of knowledge about physical principles was demonstrated by Chi, Glaser, and Rees (1981) in several different tasks. For example, experts demonstrated deeper and more complete understanding than novices of a task involving verbal description of physical principles. In solving physics problems, novices showed their lack of understanding by failing to make necessary inferences and by making incorrect inferences.

The process by which experts comprehend problems has striking parallels with the previously described process by which normal subjects comprehend text. However, unlike standard texts, sentences in problem descriptions are not explicitly related and thus lack normal coherence. Furthermore, they do not contain all the necessary information to establish coherence of the text. Hence expert problem solvers in a domain must draw extensively on their organized knowledge to provide inferences and knowledge to generate a coherent mental representation of the presented problem. In many domains experienced problem solvers often spontaneously externalize their representation of the situation described in the problem by drawing a diagram. Through an analysis of diagrams drawn by subjects with different levels of expertise, it is possible to track the shifts in the representation of different types of problems and the emergence of veridical and efficient representations in physics experts (Anzai, 1991). Valid diagrams of physics problems have been shown to be a highly effective means of representing relevant knowledge for the solution of presented problems (Larkin & Simon, 1987). Diagrams can be viewed as an externalized version of subjects' internal representation of a problem situation that would normally be stored in extended working memory. A distinct advantage of the diagram over the internal memory representation is that quantitative facts presented in the problem unknowns, which are quite arbitrary and thus lack inherent meaning, can be accurately and easily stored externally in the diagram.

More direct evidence for experts' superior ability to rapidly encode and store meaningful diagrams and other representative stimuli comes from the research that adapted Chase and Simon's (1973) paradigm in chess to different domains of expertise: bridge (Charness, 1979; Engle & Bukstel, 1978; go (Reitman, 1976); music notation (Sloboda, 1976); electronic circuit diagrams (Egan & Schwartz, 1979); computer programming (McKeithen, Reitman, Rueter, & Hirtle, 1981); dance, basketball, and field hockey (Allard & Starkes, 1991); and figure skating (Deakin & Allard, 1991). It is now clear that the superior memory of experts for briefly presented meaningful stimuli reflects increased storage in long-term memory relying on mechanisms consistent with the principles of skilled memory (Ericsson & Staszewski, 1989). Verbal reports of chess masters (de Groot, 1978) show that they are able to form an integrated representation of a chess position within even brief presentation times. This immediate representation is often by itself sufficient to allow the chess master to retrieve the best next move for that position. However, the chess masters always engage in extensive planning and evaluation of potential move sequences before committing to a specific move for a position. Frequently during this extended planning process, even grand masters uncover for the first time a chess move that they recognize as the best move. Charness (1981b) found that the depth to which a possible move sequence for a chess position was explored was closely related to the level of skill at playing chess, at least for chess players below the level of experts. Mental planning and evaluation of possible move sequences place greater demands on memory as the depth increases, and such a cognitive activity will be particularly tractable when long-term working memory is used to represent chess positions.

In support of the findings of remarkable capacities to explore chess positions mentally, it is well known that chess players at the master level can play chess games under blindfold conditions with only a minor reduction in performance and without any prior specialized practice (Holding, 1985). In the absence of a strict time constraint, there appears to be no clear limit on the depth to which a chess master can explore a position. Ericsson and Oliver (Ericsson & Staszewski, 1989) found that a candidate chess master was able to access all the information about a mentally generated chess position rapidly and accurately, and they showed that the memory representation of the chess position was consistent with the characteristics of skilled-memory theory (Chase & Ericsson, 1982; Ericsson & Staszewski, 1989). It is important to note that the chess master kept a mental representation of the complete chess position with all its chess pieces so that he would be able to explore all kinds of future chess moves and thus not be hindered by constraining semantic interpretations.

A similar relation between superior memory and ability to plan is found in the game of bridge. Charness (1989) showed that expertise in bridge was closely linked with the capacity to generate successful plans for playing the cards in the correct order.

In the task domains discussed so far, a tight association between the initial representation of the problem and the solution processes consistently emerges with extensive experience and expertise. Only with novices does retrievability of relevant information present a real difficulty.

Expert Problem Solving with Large Demands on Retrieval from Long-Term Memory

Several types of problem solving and directed thinking make extensive demands on retrieval from long-term memory. For these tasks, researchers presume that the subjects have all the necessary knowledge in long-term memory. Hence, instructions can be quite brief and primarily contain information describing the final product or answer. The critical problem in these tasks is finding organized access to stored information. In this section, we describe some research on anagram problem solving, which involves locating a single word in the vast mental lexicon. Then we discuss cognitive processes involved in retrieving information to generate more complex cognitive structures such as computer programs, texts, and solutions to complex problems in social science.

When subjects are given anagrams, they are told to rearrange a sequence of letters, for example, SCALIO, so that all letters form a word (SOCIAL). Subjects do not systematically generate the nearly 1000 possible permutations of the six letters, but rather seek out common combinations of one to three letters and, using these combinations as cues, try to elicit the word from memory. Studies of anagram problem solving with subjects instructed to think aloud (Mayzner, Tressert, & Helbock, 1964; Sargent, 1940) show that subjects generate such cues. These studies also find that subjects fairly often retrieve complete words with letters quite different from the target letters, evidence that the partial letter combinations serve as cues in the retrieval process. In some detailed analyses, Sargent (1940) showed that retrieval of words was nearly always mediated by cues consisting of letter combinations. When "good" letter cues had been generated, both good and poor anagram solvers were likely to retrieve the target word. The cognitive processes responsible for generating cues appeared to be the major difference between good and poor anagram solvers (Sargent, 1940). In sum, anagram problem solving can best be viewed as a process of generating appropriate constraints using some of the letters, which in turn are used as cues in the retrieval of words from a person's mental lexicon.

Unlike anagram problem solving, most other cognitive activities demanding extensive retrieval involve the retrieval and organization of large amounts of different kinds of knowledge. Because of vast individual differences in available knowledge, research on design of computer programs, complex social problem solving, and composition of texts has predominantly focused on structural differences in novices' and experts' organization of cognitive processes. Across these three very different task domains, remarkably consistent differences between experts and novices have been found. Not surprisingly, novices tend to start working on computer programs or their text without much, if any, planning and preparation. This is most striking for novice writers, who write their texts as a sequence of ideas generated from memory using the provided title and their most recently generated idea as cues. In a recent book, Bereiter and Scardamalia (1987) present a large body of evidence showing that novice writers produce their texts from repeated probing of memory (knowledge telling) without any prior attempt to reorganize and transform this knowledge for the reader. In contrast, all types of experts tend to spend a substantial amount of time in reviewing the problem, considering constraints, and decomposing the problem into a sequence of subproblems that can be solved independently or with minimal interaction.

Basically, solution of a complex design problem involves mapping in detail the interactions between various parts of the design product. In a computer program, for example, all modules must use the same data representation, and the calculations of certain modules must be passed to other modules. Experts in software design generate complete models of the program at increasing levels of specificity before starting to produce the solution, that is, actually writing the code for the program (Jeffries, Turner, Polson, & Atwood, 1981). According to Jeffries et al. (1981), this design process is guided by design schemes available in experts' long-term memory. In a simpler task studied by Byrne (1977), skilled subjects generated a series of three-course meals for the same guests. In generating the components of the main and second courses, subjects systematically started with the type of protein or meat, and then decided on starch and vegetables, as a sequence that would facilitate the coordination of these components. Byrne (1977) even found that subjects would frequently generate the courses out of order and start with the main and second courses to make the selection of a coordinated first course easier. Voss, Greene, Post, and Penner (1983) asked different types of experts and novices to describe how the Soviet Union could increase its agricultural production. Experts on the Soviet Union tended to begin by outlining the history of previous attempts to solve this problem and explored general solutions to increasing depths of specificity in a hierarchical manner. The hierarchical approach allowed experts to consider the general problem of capital investment and its observable implications for more specific problems, such as lack of repair parts and fertilizer. The novices, on the other hand, tended to enumerate these "surface" problems as separate solutions. Domain-specific knowledge about the Soviet Union was found to be critical, and expert political scientists without that knowledge were able to generate only the abstract issues without being able to proceed to detailed solutions. Experts in subjects outside of political science, like chemistry, tended to resemble novices in both the content and structure of their solutions.

In contrast to the experts on the Soviet Union, expert writers are frequently not experts on the topics they are writing about. Based on thinking-aloud studies of writing, Hayes and Flower (1980) demonstrated that experienced writers tend to start by generating ideas and content and spend considerable time organizing this content before beginning to write. Hence, it appears that the experienced writers construct a well-formed organization for their ideas, whereas the domain-specific experts are able to draw on preexisting cognitive structures.

The composing processes of serious poets and control subjects in an early think-aloud study by Patrick (1935) are similar in structure to those of expert and novice writers. Patrick found that the first phase of writing a poem was dominated by the generation of possible themes, ideas, and images, although she described this phase by the changes in ideas and themes and stressed the noncumulative character of subjects' thoughts. As a theme emerged, the subjects started writing lines of the poems and toward the end of their work were occupied mostly with revisions or editing.

In summary, the results from analyses of experts' cognitive processes in tasks requiring extensive retrieval tell us something about the role of memory. Through the initial planning and design activity, the expert is able to hierarchically decompose the problem or product into subproblems or subactivities (Simon, 1973). Hence, the actual production of the solution or product, which would be taxing to working memory, has been simplified and limited to a succession of single subproblems.

If subjects are to be able to focus their processing resources on a given subproblem, then the hierarchically organized solution plan has to be stored somewhere. Although few studies have tried to assess empirically where and how the solution plan is stored, several sources of evidence point to storage in long-term memory. A rich and organized knowledge structure appears to be a prerequisite for generation of a solution plan. The evidence on skilled memory previously reviewed implies that such a knowledge structure would be highly appropriate for rapid storage and efficient retrieval of information.

Other Cognitive Activities Involving Integration of Information

In several different cognitive tasks such as judgment and decision making, concept learning, and diagnosis, many pieces of information must be integrated. Because of the relatively large amount of information that needs to be considered and often stored in memory, the operation of memory in these cognitive activities is particularly interesting. As before, we consider the availability of external memory and of relevant knowledge and procedures, that is, the degree of expertise.

We first briefly discuss results from laboratory studies in which the influence of subjects' relevant knowledge was minimized. In the typical study of concept formation, subjects are asked to find the rule for distinguishing positive and negative instances. Each instance is defined by one of several possible values, such as red, blue, green, for each of a number of dimensions, such as color, shape, and size. Before the experiment starts, subjects are told what the critical dimensions are and what their possible values are. Investigators' selection of dimensions and values is quite arbitrary, and thus the subjects have no knowledge about dimensions and instances that is relevant for the concept-formation task. Because of this lack of relevant knowledge and the similarity of the instances, subjects ought to show poor memory for previously presented instances. This expectation has been supported by subsequent tests of subjects' memory of presented instances (Coltheart, 1971). Process analyses of cognitive processes in concept formation show that subjects generate a hypothesis about the rule and maintain this hypothesis in short-term memory until the task is solved or the hypothesis is disconfirmed. When the hypothesis is disconfirmed, subjects generate a new hypothesis, primarily relying on the most recently presented instances. Hence, this type of concept formation appears to use predominantly information stored in the limited capacity of short-term memory (Bourne, Goldstein & Link, 1964). The limitation of storage of information in short-term memory appears to depend on the type of stimulus material used. With more complex and meaningful materials, early investigators using complex drawings and designs (Heidbreder, 1924) as well as contemporary investigators using drawings of fantasy creatures (Brooks, 1978) have found evidence for extensive storage of exemplars in long-term memory. They have also found a wide range of mechanisms allowing subjects to discover concepts as well as to rely on analogy with stored instances to display rule-like behavior.

Laboratory research on logical reasoning has shown that college students in some tasks, even after considerable training in logic, often perform poorly with controlled abstract stimuli for which the students lack relevant knowledge and experience. Accuracy in logical reasoning is dramatically improved when the logical problems are expressed in terms from a familiar context (Griggs & Cox, 1982; Wason & Johnson-Laird, 1972). Even in simple logical tasks such as 3-term series problems, subjects used different representations to integrate the presented information (Egan & Grimes-Farrow, 1982). With more complex problems such as syllogistic reasoning, there is now considerable evidence for subjects' relying on mental models to allow them to reason about the different logical alternatives (Johnson-Laird, 1983). Reasoning by experts in their domain of expertise is normally described with reference to one or more mental models (Gentner & Stevens, 1983) that provide the knowledge and retrieval structures to allow the use of long-term working memory. There is every reason to believe that the same types of long-term working memory systems mediate the planning and evaluation of chess moves by chess experts discussed earlier and the diagnostic reasoning by medical experts discussed later in this section.

In most laboratory research on decision making, subjects have to select the best of a number of presented alternatives, which are characterized by values along a number of dimensions. Normally, all information about the alternatives is perceptually available and thus stored in external memory. The subjects are sufficiently familiar with the dimensions that they can pick out important ones and judge alternatives along each dimension with relative ease. Process analyses (Payne, 1976; Svenson, 1979) of decision making of this kind show that subjects do not integrate information about alternatives into holistic evaluations. Instead, subjects attempt to eliminate alternatives by considering their values on critical dimensions. Only when a couple of alternatives remain do subjects engage in detailed evaluation in which values on different dimensions are traded off. Subjects' systematic efforts to reduce the information that they must consider is highly consistent with the assumption that subjects are constrained to use short-term memory in decision making when they lack extensive knowledge and experience of a situation. Analyses of expert decision makers, such as financial analysts, auditors, and graduate admissions officers, have shown that experts proceed quite differently (See Camerer and Johnson [1991] for a review). The experts tend to focus on a relatively small set of important variables selected on the basis of extensive knowledge about the domain. This selected information is interrelated and interpreted in a way similar to diagnostic reasoning in medical experts.

Diagnosis, in particular medical diagnosis, is a challenging cognitive problem because a lot of presented information must be organized in such a way that the correct diagnosis can be elicited from many possibilities. Some of the early approaches to medical diagnosis (Wortman, 1966, 1972) were similar to the approach to concept formation. Medically trained subjects were given a clue or symptom and then asked to establish the disease by asking questions of the experimenter. From the subjects' thinking-aloud protocols, Wortman (1972) could determine that essentially all questions asked were motivated by hypotheses about one or several related diseases, except for general questions about the background information. Considerable additional evidence confirming a diagnosis would be gathered until the final diagnosis was given. More recent research on medical diagnosis has used patient charts. Subjects with various degrees of medical expertise think aloud about the diagnosis while reading through information about the patient statement by statement. Analysis of these think-aloud protocols has shown that some of the differences due to expertise occur because less experienced subjects have more incorrect knowledge about diseases (Feltovich, Johnson, Moller, & Swanson, 1984; Johnson, et al., 1981). It is more interesting for our understanding of memory that the major differences due to expertise were related to the experts' more reliable access to and better integration of their knowledge about diseases (Feltovich et al., 1984; Johnson et al., 1981). The medical experts were better able to integrate information about the patient and discover inconsistencies, and they could recover more easily from incorrect diagnostic hypotheses they had generated earlier. Patel and Groen (1986) have shown that after a brief review of the patient chart, the medical expert is able not only to recall the relevant information but also to give an integrated account of the underlying pathophysiology of the case. With increasing expertise, subjects appear to be able to form in long-term memory a well-integrated representation of the presented information akin to the representations found in the normal comprehension of text, discussed earlier in this article.

Summary and General Discussion

A basic assumption of human information processing is that complex cognitive processes can be described as a sequence of generated states. Any one of these states provides an appropriate characterization of the information-processing activity that has previously occurred in a task up to that point and is thus a description sufficient to serve as input for the information processing activity determining the next state. Thus, we distinguish between the massive, transient activation of many cognitive structures in the generation of a stable state, and the stable activation of a limited number of cognitive structures that constitutes a cognitive state.

This important distinction allows us to represent the horizontal dimension of thinking as a sequence of states in which each state can be described in terms of a limited number of cognitive structures. Some of the difficulties investigators have encountered in limiting the amount of information that has to be kept active in working memory could be resolved by a clearer separation of processes based on transient activation from processes leading to stable intermediate products for the next state. In our review of text comprehension, we found that studies of priming supported this distinction. Empirical data on on-line reading shows that comprehension of text can best be described as a sequence of meaningfully encoded phrases and sentences. Likewise, problem solving and other complex cognitive activities for novices as well as experts can best be described as a sequence of generated thoughts or states.

The most important aspect of conceptualizing cognitive processes as a sequence of states is the claim that each state is sufficient to characterize all preceding relevant cognitive activity. The pioneering work by Newell and Simon (1972) showed that each state could be adequately described by a small number of cognitive structures in short-term memory--well within the capacity of the seven plus or minus two chunks in short-term memory that Miller (1956) proposed. Our review of studies on problem solving, concept formation, and decision making has supported the generalizability of this claim. However, we found that these studies have been conducted in unfamiliar laboratory environments, where interference and lack of relevant knowledge and memory skills make it virtually impossible to exploit the storage of generated results and products in long-term memory. In contrast to the documented absence of storage in long-term memory in these studies, recent studies of expert performance in a wide range of domains have demonstrated rapid storage and the efficient use of extended working memory based on storage in long-term memory.

A central assumption in the information processing approach is that the complexity of the states of cognitive processes is constrained by the limited capacity of short-term memory. In turn, this limitation and is viewed as a powerful constraint on acceptable simulation models for cognitive processes in different tasks. The empirical evidence for this assumption is compelling for the restricted domain of unfamiliar tasks used in laboratory research on problem solving, concept formation, and decision making. In contrast, we have found equally compelling evidence that subjects can rapidly store long-term information about displayed stimuli as well as many aspects of their preceding cognitive processes. Furthermore, these subjects can rapidly retrieve this information when it is needed. The difference between our findings and those of classical laboratory studies is that the subjects who used long-term memory were performing tasks in highly familiar domains.

Psychologists have been reluctant to accept the possibility that intermediate products and results can be stored in long-term memory during cognitive processing. Their skepticisms is due to the apparent lack of a useful constraint on storage, the storage capacity of long-term memory having been judged to be vast and essentially unlimited. In response to those concerns, our proposal for long-term working memory establishes a set of new constraints on the type of information that can be rapidly stored in retrievable form in long-term memory. We also specify the circumstances under which this information can be so stored and retrieved.

The primary constraint for relying on long-term memory for extended working memory is the availability of an extensive body of highly organized knowledge and patterns (semantic memory) relevant to the particular domain. Studies cited in our review show that novices must deliberately generate retrieval cues to retrieve facts and knowledge from long-term memory. Experts who have extensive additional experience, rapidly and reliably access relevant facts and knowledge as part of the normal course of processing for the task. Following Chase and Ericsson's (1982) principles of skilled memory theory, we propose that meaningful encoding, or rapidly generated encoding of information in terms of pre-existing knowledge, is a prerequisite for efficient storage in long-term memory.

Once information has been rapidly encoded and stored in long-term working memory, the critical question is how this information can be efficiently retrieved when it is relevant. Unlike the limited capacity of short-term memory, the vast capacity of long-term memory requires that selective retrieval be based on relevant cues in short-term memory. Rapid retrieval from long-term memory is assumed even in the traditional models based on short-term working memory. When a large chunk is stored in short-term memory it is not assumed that all the specific information of this large chunk is directly available, but rather that a higher-level cue is available in short-term memory with which subjects can rapidly retrieve specific information from long-term memory at a rate of about 300 ms. In complex cognitive processes like text comprehension, during which an integrated representation of presented text is generated, it is reasonable to postulate that a higher-level cue in short-term memory can allow rapid retrieval from long-term working memory, speed of retrieval being estimated at around 400 ms. In other complex cognitive activities the presented information cannot be immediately integrated at its initial presentation and in these activities long-term working memory is mediated by encoded associations to retrieval structures allowing access to pieces of information with retrieval times below half a second.

In complex cognitive processes, during which presented information and intermediate products have to be stored away in long-term working memory for future use, we propose the retrieval mechanism outlined in Chase and Ericsson's (1982) skilled memory theory. At the time of the original processing of presented and generated information, subjects associate specific retrieval cues with the information stored in long-term memory working memory. By reinstating a given retrieval cue in attention, they can retrieve the associated information from long-term memory. Through the construction of large integrated structures in long-term memory and the association of retrieval cues, subjects are able to use long-term memory as an extended working memory.

We have reviewed extensive empirical evidence for storage in long-term memory during cognitive processing. First, it is a well-established fact that experts performing tasks in their domain of expertise, as well as skilled readers comprehending texts, have substantial incidental memory after completing a given task. Second, skilled readers and experts are remarkably unaffected by interruptions, even when these interruptions are designed to wipe out all the current contents of short-term memory. Third, studies of experts have shown that experts can rapidly encode and store a large amount of presented information only when meaningful and representative stimuli from their domains of expertise are used. Detailed analyses of the accessibility of the stored information support our proposal for long-term working memory. Finally, there is considerable empirical evidence for reliance on cognitive processes for planning, reasoning, and understanding, which require extensive working memory to complete.

Once it is recognized that any general theory of working memory for complex cognitive processes requires storage in long-term memory, our theoretical framework offers a parsimonious account. It is consistent with all recognized limits on human information processing and accepted mechanisms for storage and retrieval from long-term memory. The only additional assumption is that with extensive experience and practice, subjects can acquire organized memory skills that allow them to use long-term memory with storage and retrieval characteristics similar to short-term memory. Laboratory studies on the acquisition of skilled memory through extended practice (Chase & Ericsson, 1982) have demonstrated the feasibility of this assumption. The finding that the extended working memory of experts is highly domain specific strongly suggests that it reflects acquired skill.

Successful reliance of skilled memory requires considerable acquired skill. First, a subject must acquire sufficient set of retrieval cues in addition to considerable experience and practice to allow rapid encoding and storage in long-term memory. Second, the subject needs to have sufficient skill in the relevant domain of expertise to be able to anticipate future retrieval needs. At the time of original processing of the information the subject can associate it with the appropriate retrieval cues. Finally, the subject's cognitive activities need to be coordinated with the acquired memory skill in such a way that the appropriate retrieval cues are activated when the relevant information needs to be retrieved.

Our theoretical framework also provides a parsimonious account of findings which have proven to be problematic for the standard account of short-term working memory. Many investigators, and in particular Broadbent (1975), have argued that Miller's (1956) assessment of the capacity of short-term memory reflects a maximum (correct performance on 50% of trials on a <u>pure</u> memory task) and that the amount of information that can be reliably stored in short-term memory is much lower and around three or four chunks. Recently Zhang and Simon (1985) found that subjects can maintain less than three chunks in short-term memory when the chunks lack distinct verbal labels and thus cannot be maintained in short-term memory by rehearsal. Within the context of complex cognitive activities such as problem solving and decision making, the reliable working capacity of short-term memory is likely to be even lower; for example, the buffer estimates in the construction-integration theory are often as low as one or two. In a similar vein we believe that several studies of subjects' maximal short-term memory capacity in specially designed memory tasks reflect overestimates of the reliable capacity of short-term memory during normal processing, and at least in part special strategies such as active rehearsal, which are not habitually used. Assuming that long-term working memory is used more frequently than usually believed, cognitive processing is possible even with the lower estimates of reliable capacity of short-term working memory. The unimpaired text comprehension of subjects with dramatically limited short-term memory capacity supports this account. In addition, our theoretical framework provides a more parsimonious account of how memory of the comprehended text and of the processed task <u>after</u> the task is completed is related to the memory <u>during</u> the processing of the task. Even in the traditional models of short-term working memory, the old view of short-term memory as a physically distinct buffer is being replaced by a conception of short-term memory as the highly activated portion of long-term memory (Anderson, 1983; Shiffrin & Schneider, 1977).

Although our model of working memory conforms to all the basic constraints on human information processing, it asserts that subjects can acquire skill in the use of long-term memory and thereby circumvent the capacity limitations of short-term memory for specific domains and tasks. Our framework does not abolish constraints on working memory; it merely substitutes new constraints on rapid storage in and efficient retrieval from long-term working memory for the old constraints on short-term working memory. The nature of long-term working memory described in our theoretical framework raises issues very different from those studied within the short-term working memory framework. Individual differences in the capacity of working memory cannot be viewed as fundamentally fixed and unchangeable. Instead, the acquired nature of the these individual differences raises several questions. How are these individual differences acquired? How can they can be assessed for different domains and tasks? How can instructional procedures be used in remediation? Long-term working memory for tasks in a given domain of activity is an integrated part of skilled performance. It is clear that our analyses of skilled performance must probe at a deeper level of the organization of knowledge and its encoding and retrieval processes if they are to fully describe the operation of long-term working memory. Only if we are willing to describe and dissect complex cognitive skills will we ever ascertain the genuine limits of cognition and to create a theoretical framework for working memory that is adequate for the full range and complexity of cognitive processes.

Footnote

This research was supported by grant MH 15872 from the National Institute of Mental Health to Walter Kintsch and partially supported by a contract from the Army Research Institute to K. Anders Ericsson. We gratefully acknowledge the thoughtful comments on earlier drafts of this paper by Evelyn Ferstl, Randy Fletcher, Janet Grassia, and Danielle McNamara.

References

Aaronson, D. and Ferres, S. (1984). The word-by-word reading paradigm: An experimental and theoretical approach. In D. E. Kieras and M. A. Just (Eds.), New Methods in Reading Comprehension Research (pp. 31-68). Hillsdale, NJ: Erlbaum.

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: Some quantitative evidence. Journal of Experimental Psychology: Human Perception and Performance, 2, 56-70.

Afflerbach, P. P. (1990) The influence of prior knowlewdge on expert readers' main idea construction strategies. Reading Rersearch Quarterly, 25, 31-46.

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance and other domains. In K. A. Ericsson and J. Smith (Eds.), Toward a general theory of expertise: Prospects and limits (pp. 126-152). Cambridge: Cambridge University Press.

Allwood, C. M. (1982). Use of knowledge and error detection in statistical problem solving. Goteborg Psychological Reports, 12(1).

Allwood, C. M., & Montgomery, H. (1981). Knowledge and technique in statistical problem solving. European Journal of Science Education, 3, 431-450.

Anderson, J. R. (1990). Cognitive psychology and its implications . San Francisco: Freeman.

Anderson, J. R., (1987a). Methodologies for studying human knowledge Behavioral and Brain Sciences, 10, 467-505

Anderson, J. R. (1987b). Skill Acquisition: Compilation of weak-method problem solutions. Psychological Review, 94, 192-210.

Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson and J. Smith (Eds.), Toward a general theory of expertise: Prospects and limits (pp. 64-92. Cambridge: Cambridge University Press.

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. Psychological Review, 86, 124-140.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes . New York: Academic Press.

Atwood, M. E., Masson, M. E., & Polson, P. G. (1980). Further explorations with a process model for water jug problems. Memory & Cognition, 8, 182-192.

Baddeley, A. D. (1986). Working memory . New York: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 47-90). New York: Academic Press.

Baddeley, A. D., Logie, R. H., Nimmo-Smith, M. I., & Brereton, N. (1985) Components of fluid reading, Journal of Memory and Language, 24, 119-131.

Baddeley, A., Papagno, C., & Valla, G. (1988). When long-term learning depends on short-term storage. Journal of Memory and Language, 27, 586-596.

Baddeley, A., & Wilson, B. (1988). Comprehension and working memory: A single-case neuropsychological study. Journal of Memory and Learning, 27, 479-498.

Bates, E., Kintsch, W., & Fletcher, C. R. (1980). On the role of pronominalization and ellipsis in text: Some memory experiments. Journal of Experimental Psychology: Human Learning and Memory, 6, 676-691.

Bereiter, C., & Scardamalia, M. (Eds.). (1987). The psychology of written composition. Hillsdale, NJ: Erlbaum.

Berger, D. E., & Wilde, J. M. (1987). A task analysis of algebra word problems. In D. E. Berger, K. Pezdek & W. P. Banks (Eds.), Application of cognitive psychology: Problem solving, education and computing (pp. 123-137). Hillsdale, NJ: Erlbaum.

Binet, A. (1894). Psychologie des grands calculateurs et joueurs d'echecs. Paris: Libraire Hachette.

Bloom, B. S., & Broder, L. J. (Eds.). (1950). Problem-solving processes of college students. Chicago, IL: University of Chicago Press.

Bourne, L. E., Jr., Goldstein, S., & Link, W. E. (1964). Concept learning as a function of availability of previously presented information. Journal of Experimental Psychology, 67, 439-448.

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

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

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

Britton, B. K., & Gulgoz, S. (in press). Using Kintsch's model to improve instructional text: Effects of inference calls on recall and cognitive structures. Journal of Educational Psychology.

Broadbent, D. E. (1975). The magic number seven after fifteen years. In A. Kennedy and A. Wilkes (Eds.). Studies in long-term memory (pp. 3-18). London: Wiley.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch and B. B. Lloyd (Eds.). Cognition and categorization (pp. 169-211). Hillsdale, NJ: Erlbaum.

Byrne, R. (1977). Planning meals: Problem-solving on a real data-base. Cognition, 5, 287-332.

Camerer, C. F., & Johnson, E. J. (1991). The process-performance paradox in expert judgment: How can the experts know so much and predict so badly? In K. A. Ericsson and J. Smith (Eds.), Towards a general theory of expertise: Prospects and limits (pp. 195-217). Cambridge: Cambridge University Press.

Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and eror recovery in reading: A model based on eye fixations. Journal of Verbal Learning and Verbal Behavior, 20, 137-160.

Carpenter, P. A., & Just, M. A. (1977). Reading comprehension as the eyes see it. In M. A. Just, & P. A. Carpenter (Ed.), Cognitive processes in comprehension Hillsdale, NJ: Erlbaum.

Carpenter, P. A., & Just, M. A. (1981). Cognitive processes in reading: Models based on readers' eye fixations. In A. M. Lesgold, & C. A. Perfetti (Ed.), Interactive processes in reading Hillsdale, N. J.: Erlbaum.

Carver, R. P. (1990). Reading rate: A review of research and theory. San Diego: Academic Press.

Case, R. (1978). Intellectual development from birth to adulthood: A neo-Piagetian interpretation. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N. J.: Erlbaum.

Charness, N. ((1981a). Aging and skilled problem solving. Journal of Experimental Psychology: General, 110, 21-38.

Charness, N. (1976). Memory for chess positions: Resistance to interference. Journal of Experimental Psychology: Human Learning and Memory, 2, 641-653.

Charness, N. (1979). Components of skill in bridge. Canadian Journal of Psychology, 33, 1-6.

Charness, N. (1981b). Search in chess: age and skill differences. Journal of Experikmental Psychology: Human Perception and Performance, 7, 476.

Charness, N. (1989). Expertise in chess and bridge. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 183-208). Hillsdale, NJ: Erlbaum.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), Cognitive skills and their acquisition (pp. 141-189). Hillsdale, N. J.: Erlbaum.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 16, pp. 1-58). New York: Academic Press.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), Visual information processing New York: Academic Press.

Chi, M. T. H. (1976). Short-term memory limitations in children: Capacity of processing deficits? Memory & Cognition, , 455-580.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. S. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1, pp. 1-75). Hillsdale, NJ: Erlbaum.

Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), Perception and production of fluent speech (pp. 133-163). Hillsdale, NJ: Erlbaum.

Coltheart, V. (1971). Memory for stimuli and memory for hypotheses in concept identification. Journal of Experimental Psychology, 89, 102-108.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 561-584.

Deakin, J. M., & Allard, F. (1991). Skilled memory in expert figure skaters. Memory & Cognition, 19, 79-86.

de Groot, A. (1978). Thought and choice and chess. The Hague: Mouton. (Original work published 1946).

Dee-Lucas, D., Just, M. A., Carpenter, P. A., & Daneman, M. (1982). What eye fixations tell us about the time course of text integration. In R. Groner, & P. Fraisse (Ed.), Cognition and eye movements. Amsterdam: North Holland.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. Psychological Bullegin, 89, 63-100.

Dixon, P., LeFevre, J., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. Journal of Educational Psychology, 80, 465-472.

Dooling, D. J., & Lachman, R. (1971) Effects of comprehension on retention of prose. Journal of Experimental Psychoplogy, 88, 216-222.

Ebbinghaus, H. (1964; 1885). Memory: A contribution to experimental psychology (H.A. Ruger & C.E. Bussenius, Trans.). New York: Dover Publications, Inc. (Originally published, 1885).

Egan, D. E., & Grimes-Farrow, D. D. (1982). Differences in mental representations spontaneously adopted for reasoning. Memory & Cognition, 10, 297-307.

Egan, D. E., & Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. Memory and Cognition, 7, 149-158.

Engle, R. W., & Bukstel, L. H. (1978). Memory processes among bridge players of differing expertise. American Journal of Psychology, 91, 673-689.

Engle, R. W., Nations, J. K., & Cantor, J. (1990). Is "Working memory" just another name for world knowledge? Journal of Educational Psychology, 82, 799-804.

Ericsson, K. A. (1975). Instructions to verbalize as a means to study problem solving processes with the Eight Puzzle: A preliminary study (No. 458). Reports from the Department of Psychology. Stockholm: University of Stockholm.

Ericsson, K. A. (1985). Memory skill. Canadian Journal of Psychology, 39, 188-231.

Ericsson, K. A. (1988a). Analysis of memory performance in terms of memory skill. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence, (Vol. 5) (pp. 137-179).

Ericsson, K. A. (1988b). Concurrent verbal reports on text comprehension: A review. Text, 8, 295-325.

Ericsson, K. A., & Chase, W. G. (1982). Exceptional memory. American Scientist, 70, 607-615.

Ericsson, K. A., & Staszewski, J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 235-267). Hillsdale, NJ: Erlbaum.

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

Farnham-Diggory, S., & Gregg, L. (1975). Short-term memory function in young readers. Journal of Experimental Child Psychology, 19, 279-298.

Feltovich, P. J., Johnson, P. E., Moller, J. H., & Swanson, D. B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W. J. Clancey & E. H. Shortliffe (Eds.), Readings in medical artificial intelligence (pp. 275-319). Reading, MA: Addison-Wesley.

Fischer, B. & Glanzer, M. (1986). Short-term storage and the processing of cohesion during reading. The Quarterly Journal of Experimental Psychology, 38A, 431-460.

Fischer, G., Henninger, S., & Redmiles, D. (1991). Cognitive tools for locating and comprehending software objects for reuse. Thirteenth International Conference on Software Engineering, (pp. 318-328). Austin TX: ACM, IEEE`.

Fletcher, C. (1981). Short-term memory processes in text comprehension. Journal of Verbal Learning and Verbal Behavior, 20, 264-274.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. Journal of Memory and Language, 27, 235-244.

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.

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

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

Gentner, D., & Stevens, A. (1983). Mental models. Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. Journal of Memory and Language, 28, 735-755.

Gernsbacher, M. A., Varnes, K. R., & Faust, M. E. (1990). Individual differences in general comprehension skills. JOurnal of experimental Psychology: Learning, Memory, and Cognition, 16`, 430-445.

Glanzer, M., Dorfman, D., & Kaplan, B. (1981). Short-term storage in the processing of text. Journal of Verbal Learning and Verbal Behavior, 20, 656-670.

Glanzer, M., Fischer, B., & Dorfman, D. (1984). Short-term storage in reading. Journal of Verbal Learning and Verbal Behavior, 23, 467-486.

Glanzer, M., & Nolan, S. D. (1986). Memory mechanisms in text comprehension. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 275-317). New York: Academic Press`.

Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. Journal of Verbal Learning and Verbal Behavior, 13, 114-131.

Greeno, J. G. (1976). Indefinite goals in well-structured problems. Psychological Review, 83, 479-491.

Greeno, J. G., & Simon, H. A. (1988). Problem solving and reasoning. In R. C. Atkinson, R. Herrnstein, G. Lindzey & R. D. Luce (Eds.), Stevens' Handbook of experimental psychology (revised edition). New York: John Wiley & Sons.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. British Journal of Psychology, 73, 407-420.

Hanson, V. L., & Bellugi, U. (1982). On the role of sign order and morphological structure in memory for American sign language sentences. Journal of Verbal Learning and Verbal Behavior, 21, 621-633.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing precesses. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive process in writing (pp. 3-30). Hillsdale, NJ: Erlbaum.

Hayes, J. R., & Simon, H. A. (1974). Understanding written problem instructions. In L. W. Gregg (Ed.), Knowledge and cognition. Potomac, MD: Erlbaum.

Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. Castellon, D. Pisoni & G. Potts (Eds.), Cognitive theory (Vol. 2, pp. 21-41). Hillsdale, NJ: Erlbaum.

Heidbreder, E. (1924). An experimental study of thinking. Archives of Psychology, No. 73.

Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problem. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, NJ: Erlbaum.

Hjelmquist, E. (1984). Memory for conversations. Discourse Processes, 7, 321-336.

Hoffmann, J. (1927). Experimentell-psychologische Untersuchungen ueber Leseleistungen von Schulkindern. Archiv fuer die Gesamte Psychologie, 58, 325-388.

Holding, D. H. (1985). The psychology of chess skill. Hillsdale, NJ: Erlbaum.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. The Psychology of Learning and Motivation, 19, 59-87.

Hörmann, H. (1976). Meinen und Verstehen . Frankfurt: Suhrkamp.

Hunt, E. (1978). Mechanics of verbal ability. Psychological Review, 85, 109-130.

Huttenlocher, J., & Burke, D. (1976). Why does memory span increase with age? Journal of Verbal Learning and Verbal Behavior, 8, 1-31.

Hyde, T. S., & Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 12, 471-480.

Jackson, M. D. & McClelland, J. L. (1979) Processing determinants of reading speed. Journal of Experimental Psychology: General, 108, 151-181.

Jarvella, R. J. (1971). Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behavior, 10, 409-416.

Jeffries, R., Turner, A. A., Polson, P. G., & Atwood, M. E. (1981). The processes involved in designing software. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Johnson, P. E., Duran, A. A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. J., & Swanson, D. B. (1981). Expertise and error in diagnostic reasoning. Cognitive Science, 5, 235-283.

Johnson-Laird, P. N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness. Cambridge, MA: Harvard University Press.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.

Karat, J. A. (1982). Model of problem solving with incomplete constraint knowledge. Cognitive Psychology, 14, 538-559.

Keenan, J. M., MacWhinney, B., & Mayhew, D. (1977). Pragmatics in memory: A study of natural conversation. Journal of Verbal Learning and Verbal Behavior, 16, 549-560.

Kieras, D. E. (1984). A method for comparing a simulation to reading time data. In D. E. Kieras and M. A. Just (Eds.), Methods in Reading Comprehension Research (pp. 299-325). Hillsdale, NJ: Lawrence Erlbaum.

Kilpatrick, J. (1968). Analyzing the solution of word problems in mathematics: An exploratory study (Doctoral dissertation, Stanford University, 1967). Dissertation Abstracts International, 28, 4380-A. (University Microfilms No. 68-6442).

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

Kintsch, W. (1982). Memory for text. In A. Flammer, & W. Kintsch (Ed.), Discourse processing Amsterdam: North Holland.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. Psychological Review, 95, 163-182.

Kintsch, W. (in press a). Kognitionspsychologische Modelle des Textverstehens: Literarische Texte. In K. Reusser, & M. Reusser (Ed.), Verstehen lernen - Verstehen lehren Bern: Hans Huber.

Kintsch, W. (in press b). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Ed.), From learning processes to cognitive processes: Essays in honor of William K. Estes Hillsdale, N.J.: Erlbaum.

Kintsch, W., & Bates, E. (1977). Recognition memory for statements from a classroom lecture. Journal of Experimental Psychology: Human Learning and Memory, 3, 150-159.

Kintsch, W., & Keenan, J. M. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 5, 257-279.

Kintsch, W., & Mannes, S. M. (1987). Generating scripts from memory. In E. van der Meer and J. Hoffman (Eds.). Knowledge aided information processing. Amsterdam: Elsevier.

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

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilsson (Ed.), Perspectives of memory research (pp. 325-366). Hillsdale, N. J.: Erlbaum.

Kintsch, W., & Welsch, D. (in press). The construction-integration model: A framework for studying memory for text. In W. E. Hockley, & S. Lewandowsky (Ed.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock Hillsdalle, N. J.: Erlbaum.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. Journal of Memory and Language, 29, 133-159.

Ladefoged, P, & Broadbent, D. E. (1960) Perception of sequences in auditory events, Quarterly Journal of Experimental Psychology, 12, 162-170.

Lane, D. M., & Robertson, L. (1979). The generality of levels of processing hypothesis: An application to memory for chess positions. Memory and Cognition, 7, 253-256.

Larkin, J., McDermott, J., Simon, O. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

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

Levin, H., & Addis, A. B. (1979). The eye-voice span. Cambridge, MA: The MIT Press.

Levine, M. (1966). Hypothesis behavior by humans during discrimination learning. Journal of Experimental Psychology, 71, 331-338.

Lucas, J. F. (1972). An exploratory study on the diagnostic teaching of heuristic problem-solving strategies in calculus (Doctoral dissertation, University of Wisconsin, 1972). Dissertation Abstracts International, 32, 6828A. (University Microfilms No. 72-15368).

Luria, A. R. (1968). The mind of a mnemonist. New York: Avon.

Marr, D. (1982). Vision . San Francisco: Freeman.

Masson, M. E. J. (1984). Memory for the surface structure of sentences: Remembering with and without awareness. Journal of Verbal Learning and Verbal Behavior, 23, 579-592.

Masson, M. E. J., & Miller, J. A. (1983). Working memory and individual differences in comprehension and memory of text. Journal of Educational Psychology, 75, 314-318.

Mathews, A., & Chodorow, M. S. (1988). Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location, and depth of embedding. Journal of Memory and Language, 27, 245-260.

Mayzner, M. S., Tresselt, M. E., & Helbock, H. (1964). An exploratory study of mediational responses in anagram problem solving. Journal of Psychology, 57, 263-274.

McDonald, J. L., & Carpenter, P. A. (1981). Simultaneous translation: Idiom interpretation and parsing heuristics. Journal of Verbal Learning and Verbal Behavior, 20, 321-247.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. Cognitive Psychology, 13, 307-325.

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

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

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

Nathan, M. J., Kintsch, W., & Young, E. (in press). A theory of word algebra problem comprehension and its implications for the design of learning environments. Cognition and Instruction.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automization of skills. In J. R. Anderson (Ed.), Cognitive skills and their acquisition (pp. 57-84). Hillsdale, NJ: Erlbaum.

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

Norman, D. A., & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. Cognitive Psychology, 11, 107-123.

Oakhill, J. (1982) Constructive processes in skilled and less skilled readers. British Journal of Psychology, 73, 13-20.

Paige, J. M., & Simon, H. A. (1966). Cognitive processes in solving algebra word problems. In B. Kleinmuntz (Ed.), Problem solving. New York: Wiley.

Patel, V. L., & Groen, G. L. (1986). Knowledge based solution strategies in medical reasoning. Cognitive Science, 10, 91-116.

Patrick, C. (1935). Creative thought in poets. Archives of Psychology, 178.

Payne, J. W. (1976). Task complexity and continient processing in decisionmaking: An informational search and protocol analysis. Organizational Behavior and Human Performance, 16, 366-387.

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

Perfetti, C. A., & Lesgold, M. A. (1977). Discourse comprehension and sources of individual differences. In M. A. Just, & P. A. Carpenter (Ed.), Cognitive processes in comprehension Hillsdale, N. J.: Erlbaum.

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

Potter, M. C. (1983). Representational buffers: The eye-mind hypothesis in picture perception, reading, and visual search. In K. Rayner (Ed.), Eye movements in reading: Perceptual and language processes (pp. 413-437). New York: Academic Press.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. Journal of Experimental Psychology, 81, 10-15.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. Journal of Memory and Learning, 29, 633-654.

Raajmaker, J. G., & Shiffrin, R. M. (1981). Search of associative memory. Psychological Review, 88, 93-134.

Raphael, T. E., & McKinney, J. (1983). An examination of fifth- and eighth-grade children's question-answering behavior: An instructional study in meta cognition. Journal of Reading Behavior, 15, 67-86.

Recht, D. R., Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. Journal of Educational Psychology, 80, 16-20

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. Cognitive Psychology, 6, 436-450.

Reiser, B. J. (1986). The encoding and retrieval of memories of real-world experiences. In J. A. Galambos, R. P. Abelson & J. B. Black (Eds.). Knowledge structures (pp. 71-102). Hillsdale, NJ: Erlbaum.

Reitman, J. (1976). Skilled perception in go: Deducing memory structures from inter-response times. Cognitive Psychology, 8, 336-356.

Rizzo, N. D. (1939). Studies in visual and auditory memory span with specific reference to reading disability. Journal of Experimental Education, 8, 208-244.

Robinson, C. S., & Hayes, J. R. (1978). Making inferences about relevance in understanding problems. In R. Revlin & R. E. Mayer (Eds.), Human reasoning (pp. 195-206). Washington, D.C.: Winston/Wiley.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. Cognitive Psychology, 16, 371-416.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 2, 437-442.

Sargent, S. S. (1940). Thinking processes at various levels of difficulty. Archives of Psychology, 249.

Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. Cognitive Psychology, 3, 552-631.

Schank, R. C. (1982). Dynamic memory: A theory of reminding and learning in computers and people. New York: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding . Hillsdale, N. J.: Erlbaum.

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.

Schneider, W., Körkel, J., & Wienert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. Journal of Educational Psychology, 81, 306-312.

Schoenfeld, A. H. (1979). Explicit heuristic training as a variable in problem-solving performance. Journal of Research in Mathematics Education, 10, 173-187.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowsky, M. (1982). Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, 14, 489-537.

Seitz, M. R. & Weber, B. A. (1974) Effects of response requirements on the location of clicks superimposed on sentences, Memory & Cognition, 2, 43-46.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Silver, E. A. (1981). Recall of mathematical information: Solving related problems. Journal for Research and Mathematics Education, 12, 54-64.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Erlbaum.

Simon, H. A. (1973). The structure of ill structured problems. Artificial Intelligence, 4, 181-201.

Simon, H. A. (1974). How big is a chunk? Science, 183, 482-488.

Simon, H. A. (1979). Models of thought. New Haven: Yale University Press.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. Cognitive Psychology, 8, 165-190.

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1991) Individual differences in bridging inference processes. Manuscript submitted.

Sloboda, J. (1976). Visual perception of musical notation: Registering pitch symbols in memory. Quarterly Journal of Experimental Psychology, 28, 1-16.

Spillich, G. J. (1983). Life-span components of text processing: Structural and procedural differences. Journal of Verbal Learning and Verbal Behavior, 22, 231-244.

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

Staszewski, J. J. (1988a). Skilled memory and expert mental calculation. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.). The nature of expertise (pp. 71-128). Hillsdale, NJ: Erlbaum.

Staszewski, J. J. (1988b). The psychological reality of retrieval structures: An investigation of expert knowledge (Doctoral dissertation, Cornell University, 1987). Dissertation Abstracts International, 48, 2126B.

Svenson, O. (1979). Process description of decision making. Organizational Behavior and Human Performance, 23, 86-112.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 18, 523-534.

Tabossi, P. (1988) Effects of context on the immediate interpretation of unambiguous nouns. Journal of Experimental Psychology: :Learning, Memory, and Cognition, 14, 153-162

Till, R. E., Mross, E. R., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. Memory & Cognition, 16, 283-298.

van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension . New York: Academic Press.

Vonk, W. (1985) The immediacy of inferences in the understanding of pronouns. Inferences in text processing. G. Rickheit & H. Strohner (Eds.) Amsterdam: North Holland. Pp 205-218.

Voss, J. F., Greene, T. R., Post, T. A., & Penner, B. C. (1983). Problem-solving skill in the social sciences. The Psychology of Learning and Motivation, 17, 165-213.

Walker, W. H., & Kintsch, W. (1985). Automatic and strategic aspects of knowledge retrieval. Cognitive Science, 9, 261-283.

Wason, P. C., & Johnson-Laird, P. N. (1972). Psychology of reasoning: Structure and content. Cambridge, MA: Harvard University Press.

Webb, N. L. (1975). An exploration of mathematical problem-solving processes (Doctoral dissertation, Standford University, 1975). Dissertation Abstracts International, 36, 2689A. (University Microfilms No. 75-25625).

Whitten, W. B., & J. M. Leonard (1981). Directed search through autobiographical memory. Memory and Cognition, 9, 566-579.

Williams, D. M., & Hollan, J. D. (1981). The process of retrieval from very long-term memory. Cognitive Science, 5(2), 87-119.

Williams, D. M., & Santos-Williams, S. (1980). Method for exploring retrieval processes using verbal protocols. In R. S. Nickerson (Ed.), Attention and performance (Vol. 8). Hillsdale, NJ: Erlbaum.

Wortman, P. M. (1972). Medical diagnosis: An information approach. Computers in Biomedical Research, 3, 315-328.

Wortman, P. M. (1966). Representation and strategy in diagnostic problem solving. Human Factors, 8, 48-53.

Yates, F. A. (1966). The art of memory . Chicago: University of Chicago Press.

Yu, B., Zhang, W., Jing, Q, Peng, R., Zhang, G., & Simon, H. A. (1985). STM capacity for Chinese and English language material. Memory & Cognition, 13, 202-207.

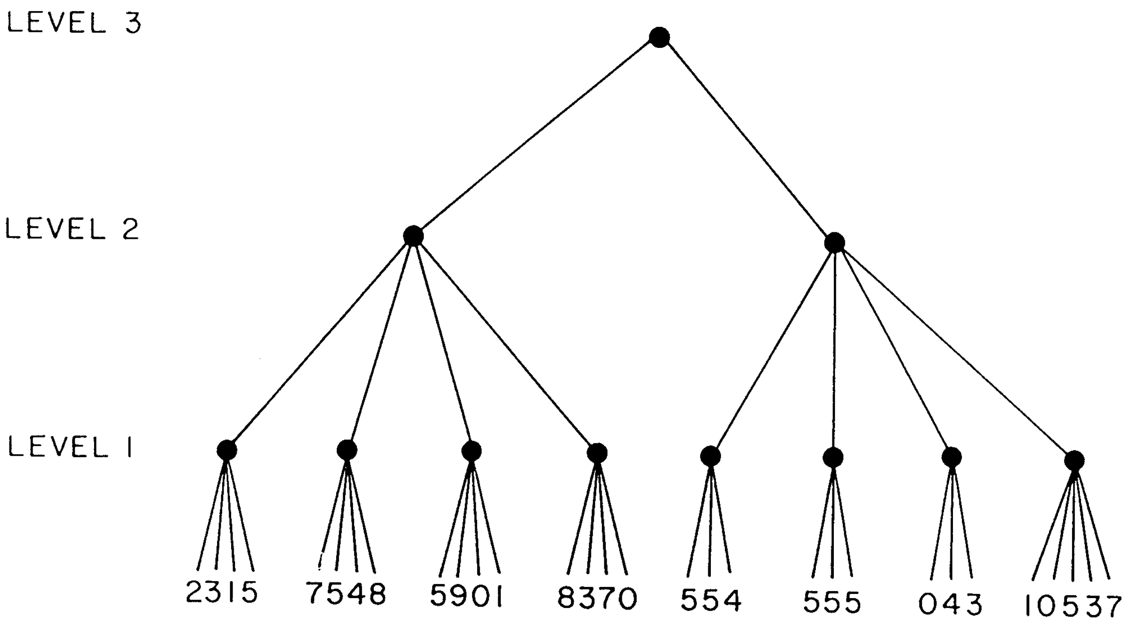
Zhang, W., & Simon, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. Memory & Cognition, 13, 193-201.

Zimny, S. T. (1987). Recognition memory for sentences from discourse. Unpublished dissertation, University of Colorado, Boulder.

List of Figures

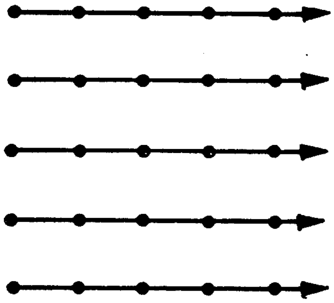
Figure 12. Working Memory, long-term memory, and the short-term buffer in text processing according to the CI model.

Figure 13. Examples of instructions and perceptually available instructions.

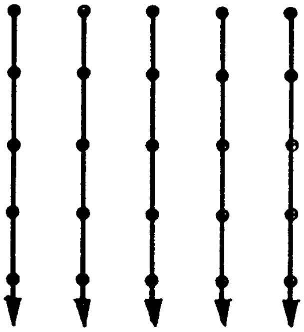

LEVEL 3
LEVEL 2
LEVEL 1
2315
7548
5901
8370
554
555
043
10537

4	7	1	0	2
3	0	4	3	6
2	1	1	4	8
8	7	4	2	9
1	5	2	7	9

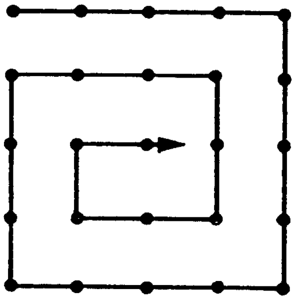
ROWS



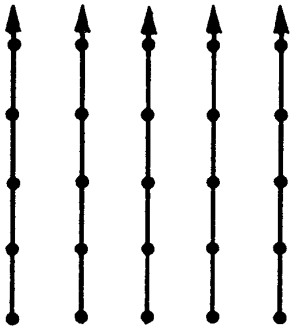
COLUMNS



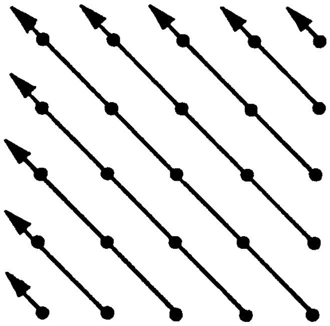
SPIRAL



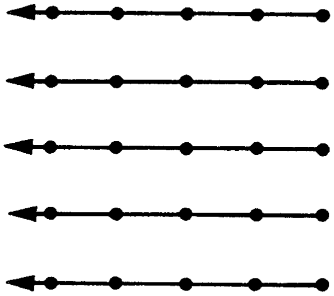
UPWARD COLUMNS

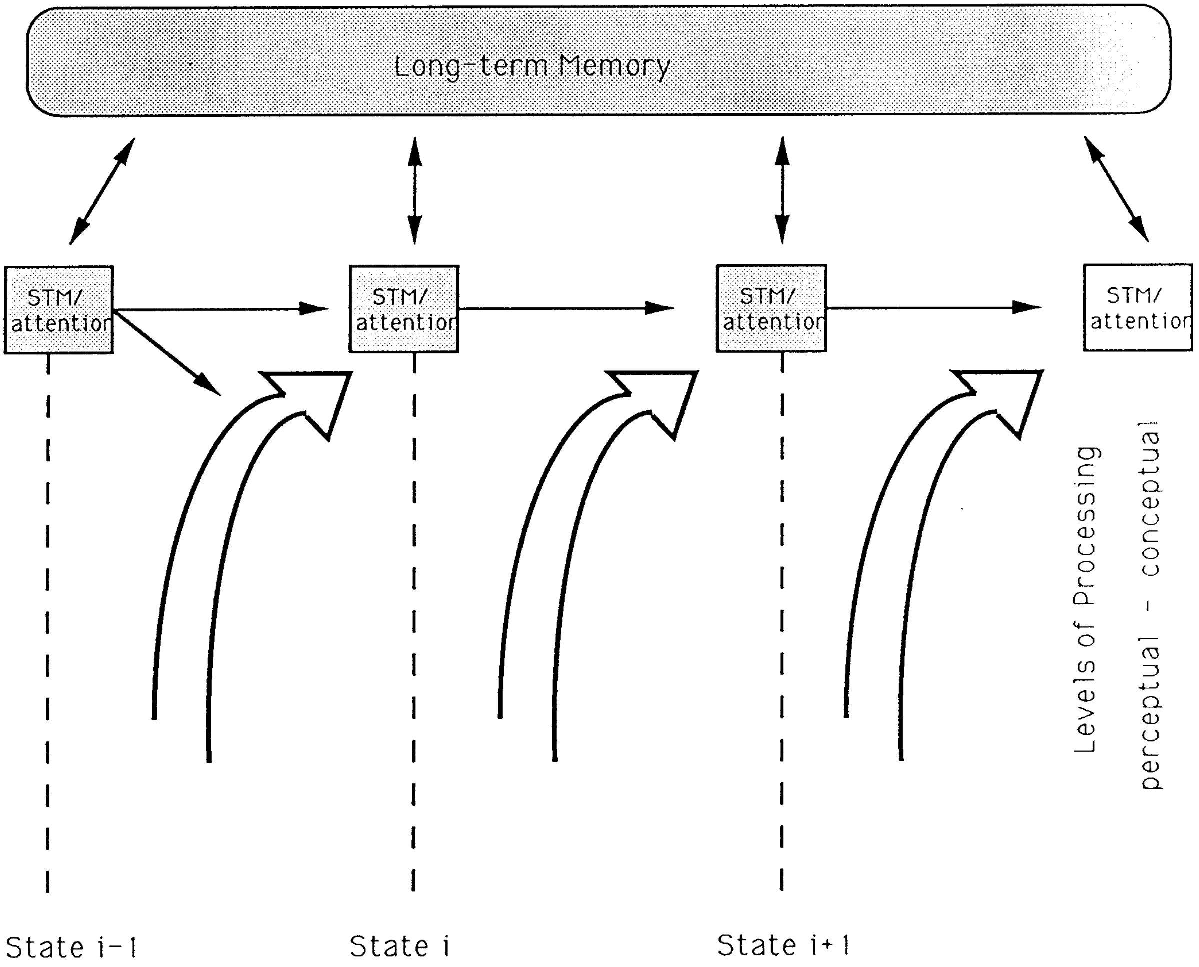


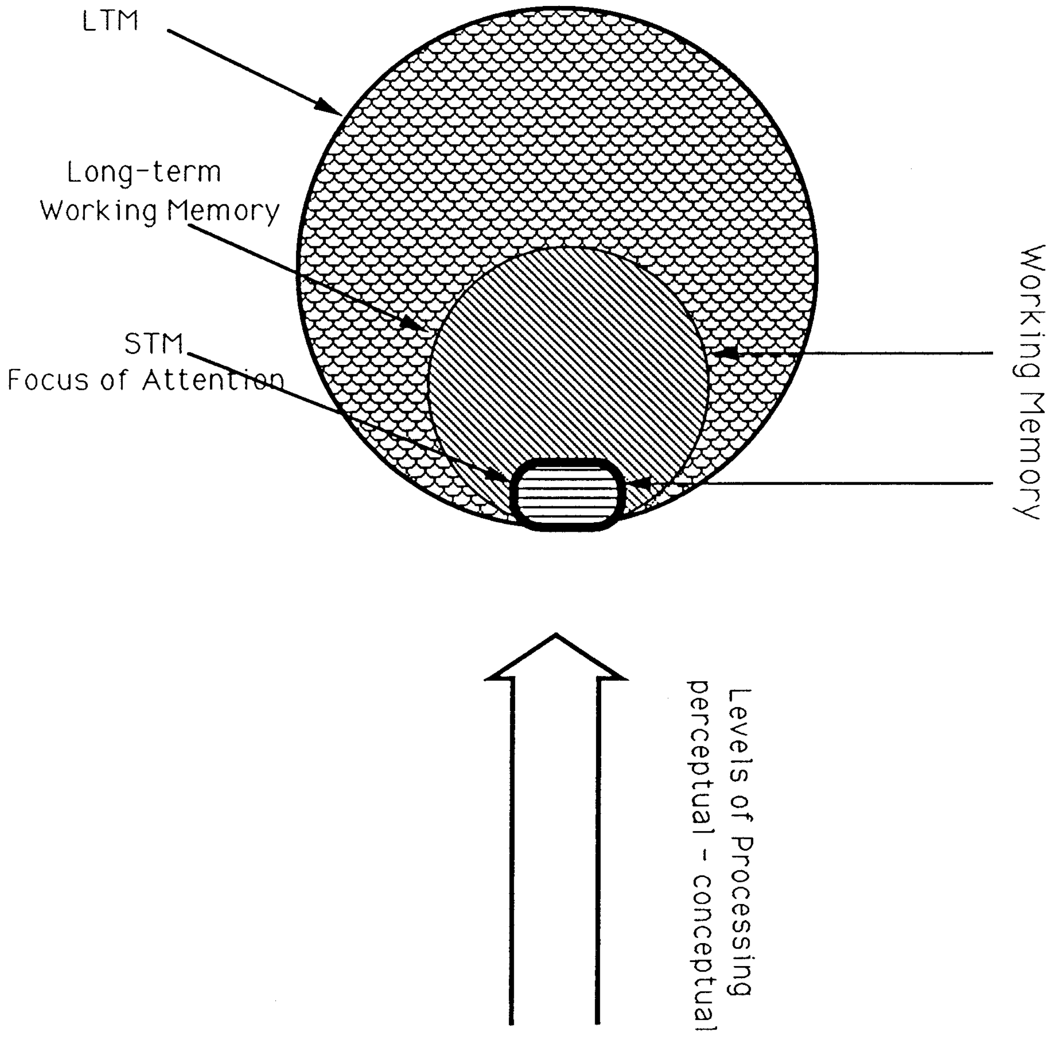
DIAGONALS

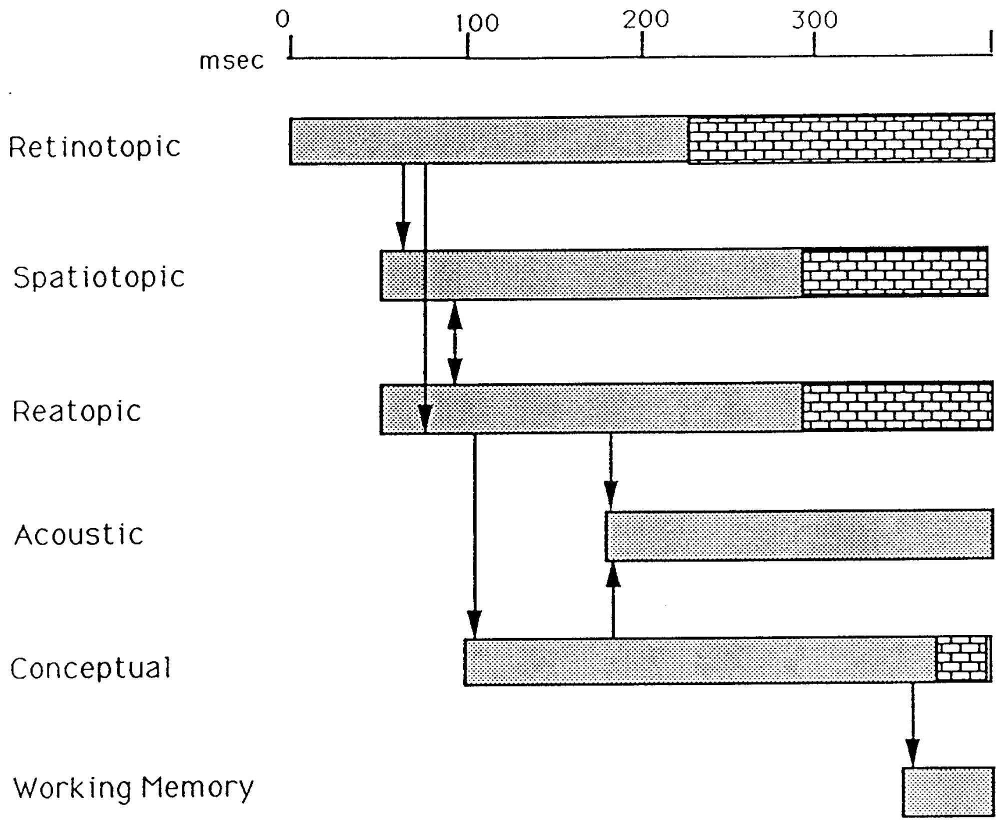


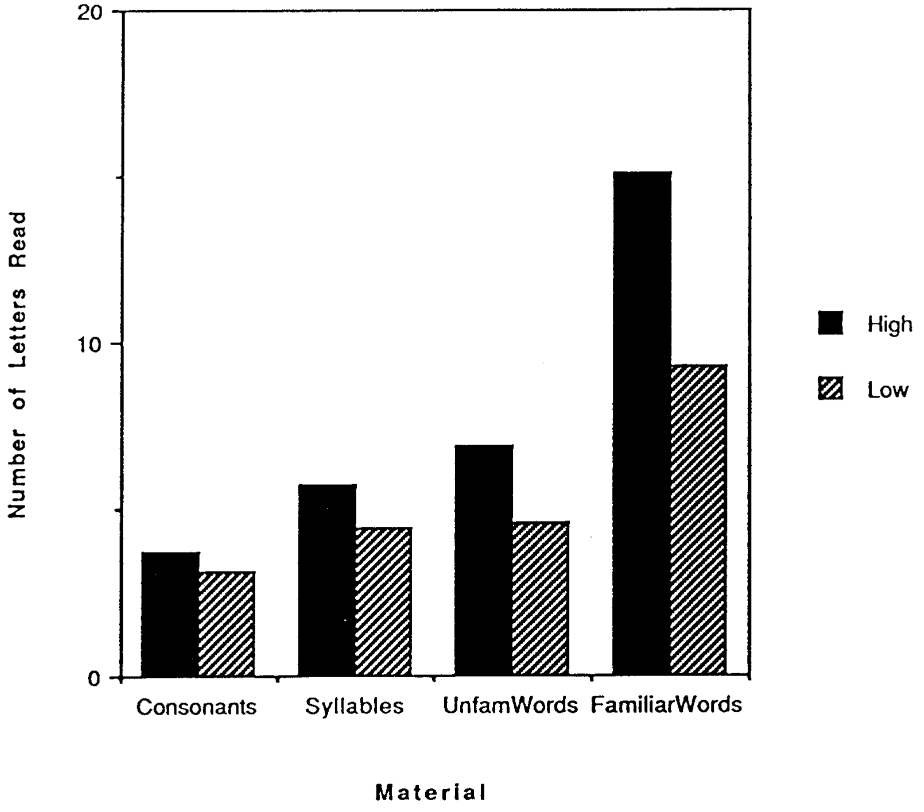
BACKWARD ROWS

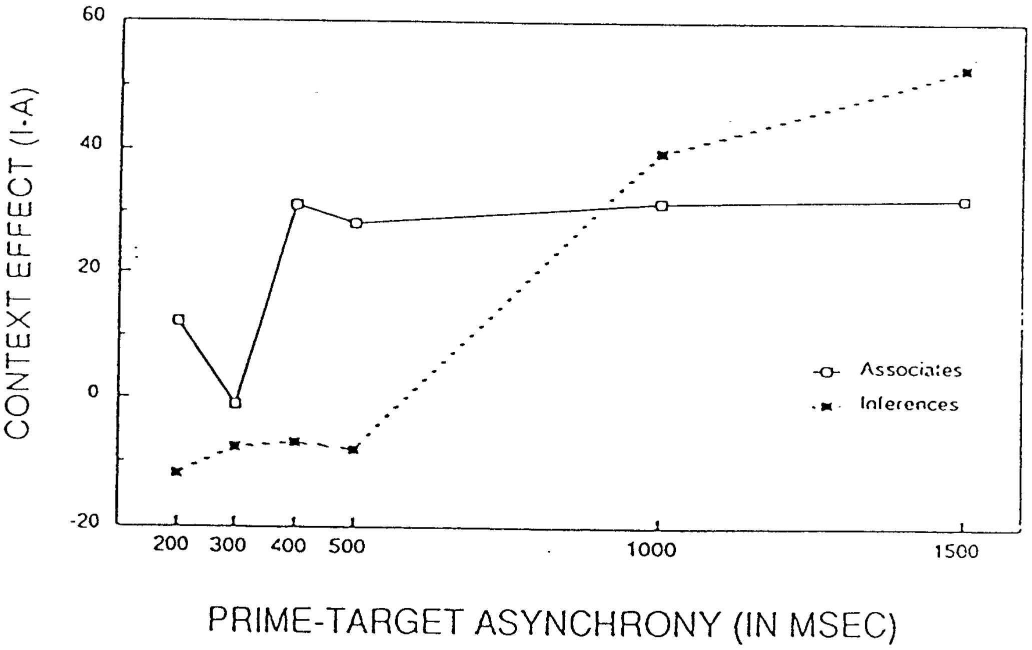


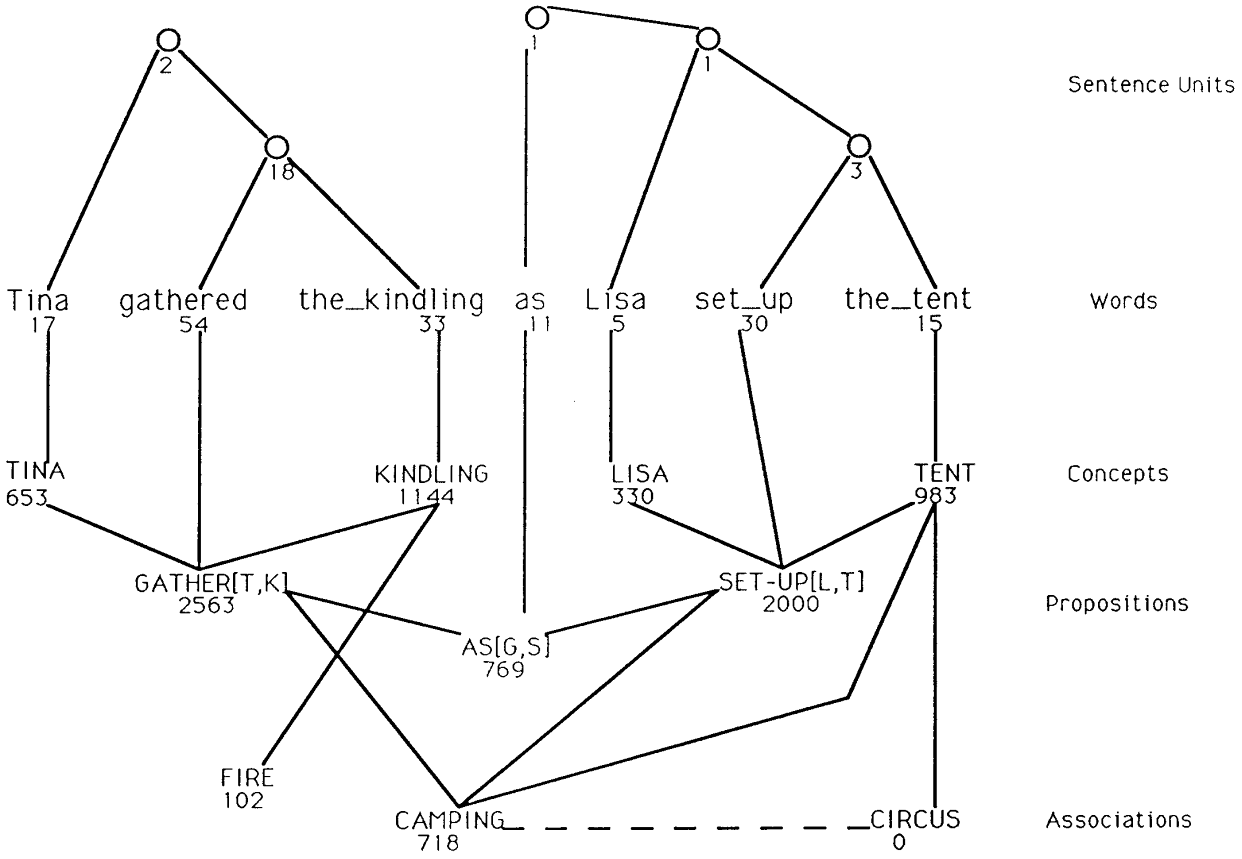

Long-term Memory
STM/ attention
STM/ attention
STM/ attention
STM/ attention
Levels of Processing
perceptual - conceptual
State i-1
State i
State i+1


LTM
Long-term
Working Memory
STM
Focus of Attention
Working Memory
Levels of Processing
perceptual - conceptual

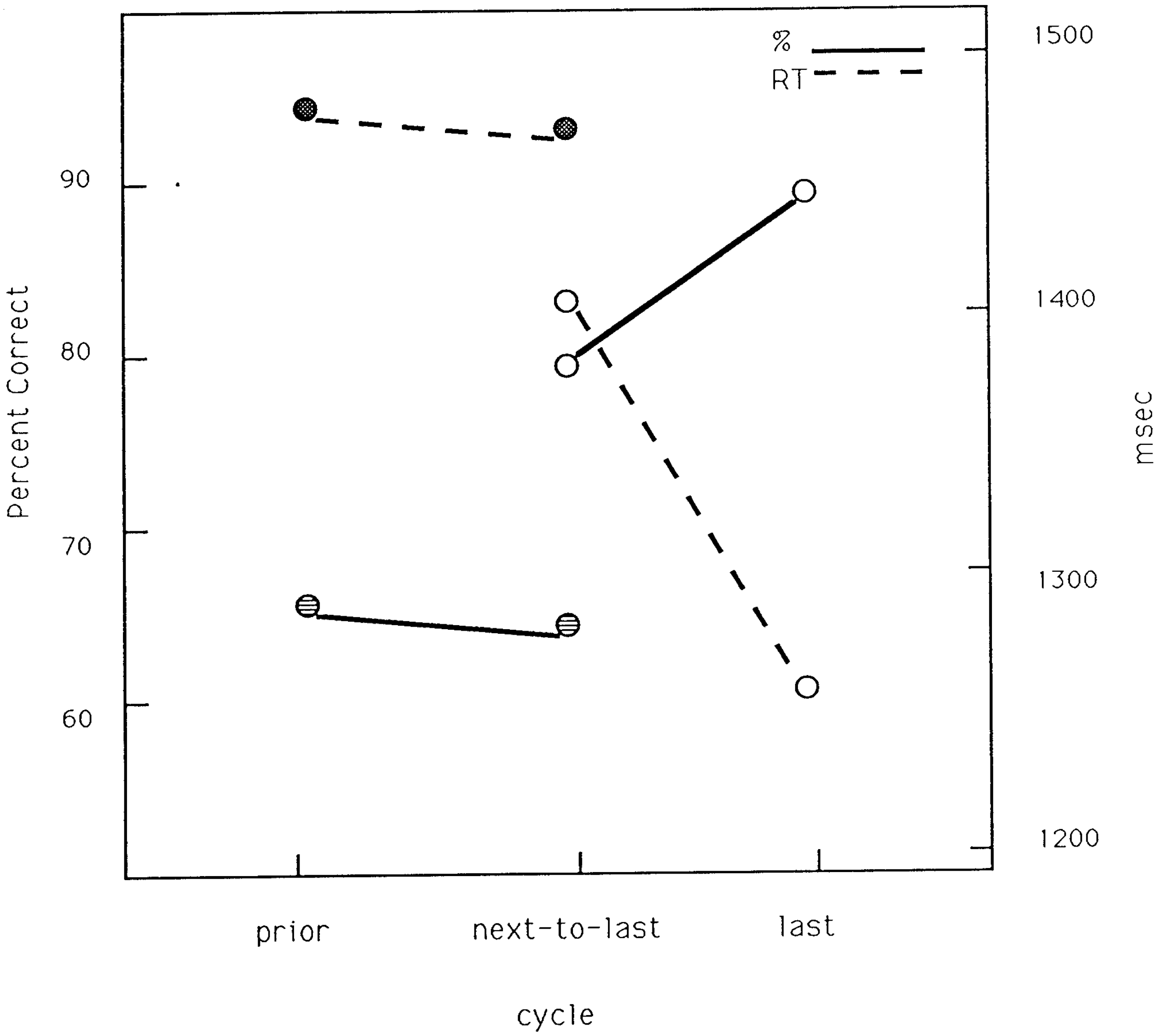

0
100
200
300
msec
Retinotopic
Spatiotopic
Reatopic
Acoustic
Conceptual
Working Memory

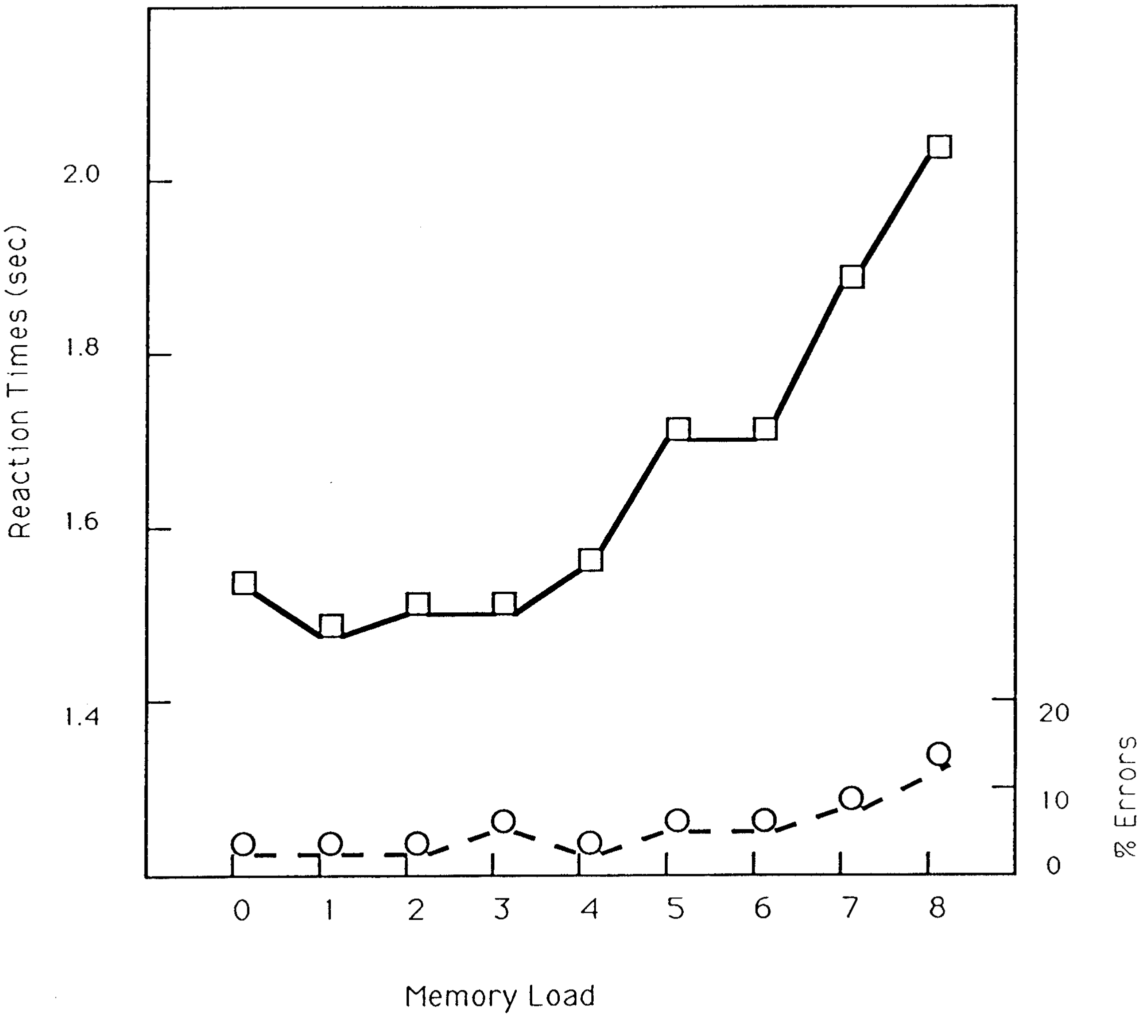

20
10
0
Number of Letters Read
Consonants
Syllables
UnfamWords
FamiliarWords
Material
■ High
▨ Low


60
40
20
0
-20
CONTEXT EFFECT (I-A)
200
300
400
500
1000
1500
PRIME-TARGET ASYNCHRONY (IN MSEC)
Associates
Inferences

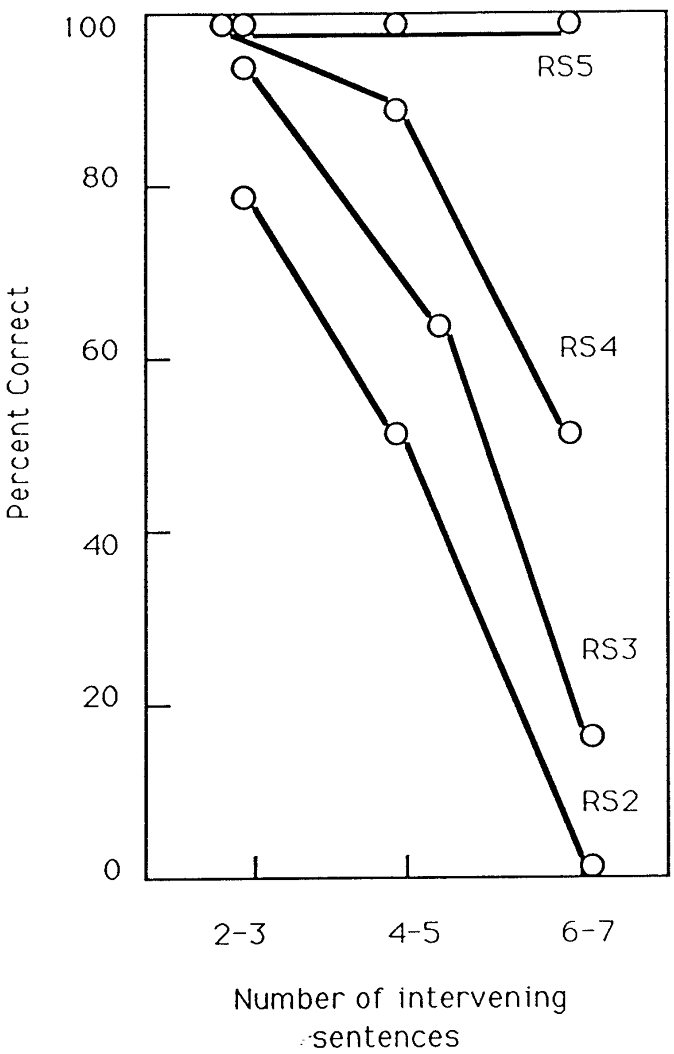


Sentence Units
2
18
1
1
3
Words
Tina 17
gathered 54
the_kindling 33
as 11
Lisa 5
set_up 30
the_tent 15
Concepts
TINA 653
KINDLING 1144
LISA 330
TENT 983
Propositions
GATHER[T,K] 2563
AS[G,S] 769
SET-UP[L,T] 2000
Associations
FIRE 102
CAMPING 718
CIRCUS 0

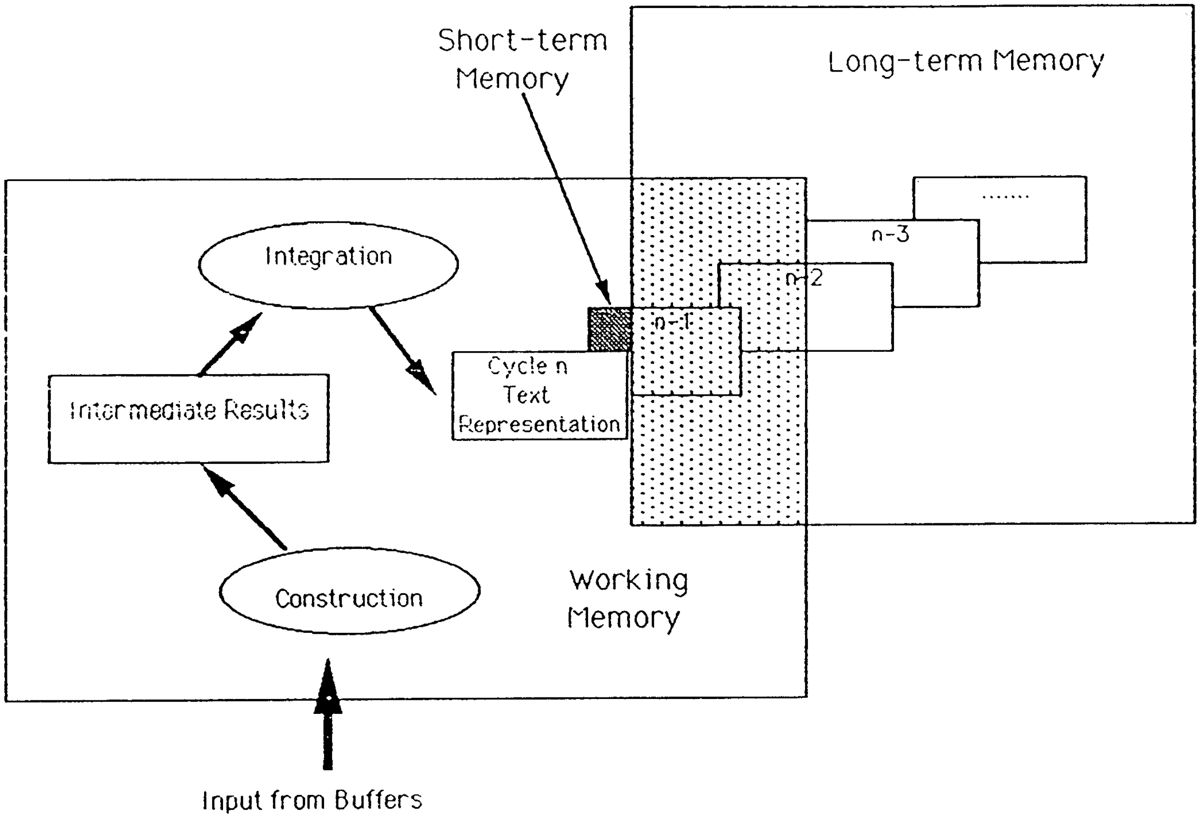

%
RT
90
80
70
60
Percent Correct
1500
1400
1300
1200
msec
prior
next-to-last
last
cycle



Reaction Times (sec)
2.0
1.8
1.6
1.4
20
10
0
% Errors
0
1
2
3
4
5
6
7
8
Memory Load



100
80
60
40
20
0
Percent Correct
RS5
RS4
RS3
RS2
2-3
4-5
6-7
Number of intervening sentences


Short-term Memory
Long-term Memory
Integration
....
n-3
n-2
n-1
Cycle n Text Representation
Intermediate Results
Construction
Working Memory
Input from Buffers

8 - Puzzle

Instruction :

Move the tiles in the puzzle such that the initial configuration on the left is transformed into the goal configurationshown on the right.

4	5	8
1	2	3
7	6	

→

1	2	3
4	5	6
7	8	

During problem solving :

1	4	5
7	2	8
6	3	

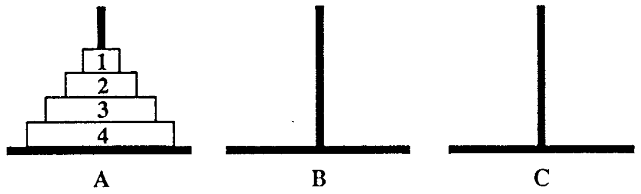
Comment :

The constraints for moving tiles are physical and cannot be violated without breaking the frame of the puzzle, which holds the tiles together.

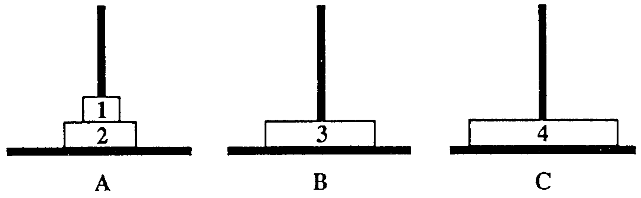
Tower of Hanoi

Instruction :

Move all discs to the peg on the right, so they will be in the same order. You may move only one disc at a time. You must not place a larger disc on a smaller. A disc moved from one peg must be placed on one of the other pegs.



During problem solving :



Comment :

Experimenter reminds the subject about rules and monitors the legality of moves.