Psychological Review
1992, Vol. 99, No. 3, 440–466

0033-295X/92/$3.00

Inference During Reading

Gail McKoon and Roger Ratcliff
Northwestern University

Most current theories of text processing assume a constructionist view of inference processing. In this article, an alternative view is proposed, labeled the *minimalist hypothesis.* According to this hypothesis, the only inferences that are encoded automatically during reading are those that are based on easily available information, either from explicit statements in the text or from general knowledge, and those that are required to make statements in the text locally coherent. The minimalist hypothesis is shown to be supported by previous research and by the results of several new experiments. It is also argued that automatically encoded minimalist inferences provide the basic representation of textual information from which more goal-directed, purposeful inferences are constructed.

In reading, comprehension processes are generally assumed to combine information from two sources: explicit statements from the text being read and general knowledge already known to the reader. Interactions of information from these two sources produce the representation of a text that is encoded into memory. The issue addressed in this article is the extent to which these interactions lead to the encoding of inferences. We claim that there is only minimal automatic processing of inferences during reading. Our hypothesis is that readers do not automatically construct inferences to fully represent the situation described by a text. In the absence of specific, goal-directed strategic processes, inferences of only two kinds are constructed: those that establish locally coherent representations of the parts of a text that are processed concurrently and those that rely on information that is quickly and easily available. This minimalist claim is supported in this article with several new experiments and with conclusions drawn from a review of previous research.

For different readers, minimalist processing with little strategic processing will occur in different situations. For some readers, it might be a rare occurrence; for others, it might happen in such situations as reading a magazine on an airplane, reading the newspaper through the morning fog over breakfast, or reading texts in a psychology experiment. However, more often than not, readers do have specific goals, especially when learning new information from texts, and so they often engage in strategic processes designed to achieve those goals. The minimalist claim for these situations is that minimal inferences provide the database for more strategic processes. They provide the database for strategic inferences that are constructed during reading, and they provide a minimalist representation of a text in memory from which strategic inferences can be constructed by retrieval operations.

The minimalist position is presented as an hypothesis from which to work toward explicit processing models. The hypothesis distinguishes between those inferences that are labeled *automatic* and those that are labeled *strategic;* however, this distinction is not always clear cut. In situations where a reader adopts special strategies, some strategic inferences may be easy to construct, perhaps nearly as easy as minimal inferences. Some strategic inferences may also be obligatory, in the sense that the text cannot be completely understood without them (Gerrig, 1986). It is our hope that an understanding of what information is provided quickly and automatically will provide the basis for an understanding of which effortful strategic and goal-based processes are relatively easy to construct and which more difficult. In fact, if a strict automatic–strategic demarcation is not eventually tenable, then the product of the minimalist program will be a set of results that label inferences in terms of speed of availability, ease of processing, probability of occurrence, and dependence on contextual environment. These results are critical in the development of processing models.

For present purposes, an *inference* is defined as any piece of information that is not explicitly stated in a text. This definition includes relatively simple inferences as well as complex, elaborative inferences and inferences that add new concepts to a text as well as those that connect pieces of the text. For example, by this definition it would be an inference to encode the relation between a pronoun and its referent or to encode two instances of the same word as referring to the same concept. It would also be an inference to compute *2* as the referent of *the number that is four less than the product of three times two* or to combine the clues of a mystery novel to give the murderer. Defining *inference* this broadly emphasizes the different degrees of processing that are required to produce different inferences. Some inferences

Experiment 1 was prompted in part by discussions with Al Collins and Ed Smith. This research was supported by National Science Foundation (NSF) Grant 85-16350 and Air Force Office of Scientific Research Grant 90-0246 (jointly funded by NSF) to Gail McKoon and by NSF Grant 85-10361 and National Institute of Mental Health Grants HD MH44640 and MH00871 to Roger Ratcliff. We thank Tom Trabasso for stimulating discussions about Experiment 3 and for comments that led to Experiment 4. We also thank Richard Gerrig, Art Glenberg, and Jan Keenan for their many helpful comments and Art Glenberg for the materials used in Experiment 5.

Correspondence concerning this article should be addressed to Gail McKoon, Psychology Department, Northwestern University, Evanston, Illinois 60208.

seem to be made automatically, without awareness. Others seem to involve conscious, problem-solving types of processing.

The automatic inferences that are the focus of this article are assumed to be supported by information that is quickly and easily available, and this kind of information is assumed to come from one of two sources: well-known information from general knowledge and explicit information from the text being read. Inferences based on general knowledge have been demonstrated in the encoding of such inferences as elaborations about "what will happen next" in a story *if* what will happen next is very predictable, the encoding of inferences about aspects of the meanings of words *if* they are highly typical aspects, the encoding of inferences about instances of categories *if* the instances are highly typical, and so on. For inferences based on explicit textual information, the information may be in short-term memory or it may be easily retrievable from the long-term memory representation of the text that is under construction.

Inferences based on explicit textual information are used to establish local coherence for a text. These inferences include connections among instances of the same concept, pronominal reference, and perhaps causal relations. *Local coherence* is defined for those propositions of a text that are in working memory at the same time; in other words, propositions that are no farther apart in the text than one or two sentences. Many of the inferences that establish local coherence are based on information that is easily available because it is in short-term memory. Other local inferences, such as the relation between *the dog* and *the collie,* are based on combinations of explicitly stated information and well-known general knowledge. In either case, inference processes are assumed to proceed automatically. Only when neither explicit short-term memory information nor general knowledge leads to a coherent local representation of a text are other processes, perhaps strategic, problem-solving types of processes, engaged to provide local coherence.

According to the minimalist position, only the two classes of inferences, those based on easily available information and those required for local coherence, are encoded during reading, unless a reader adopts special goals or strategies. Automatically processed inferences are the main focus of this article for two reasons. First, they represent the most controversial point of debate between advocates of a minimalist position and advocates of a more constructionist view of text processing. There are many potential inferences that would be automatically generated during reading according to constructionist theories but not according to a minimalist view.

Second, although much of reading may have as its goal the generation of strategic inferences (e.g., in education, problem solving, planning, or decision making), these inferences must depend on the information automatically provided by a text. Automatic inferences are those that are encoded in the absence of special goals or strategies on the part of the reader, and they are constructed in the first few hundred milliseconds of processing. They therefore merit attention because they form the basic representation of a text from which other, more purposeful, inferences are constructed. In terms of theory development, our aim is to understand what kinds of information are quickly and easily available. Such an understanding is required to build processing accounts of the construction of automatic inferences. In turn, representational and processing models for automatic encoding would optimally serve as the starting point for explanations of more strategic encoding processes.

It is interesting to note the history of our approach to this minimalist position. About 12 years ago, we began experiments (prompted by discussions with Ed Smith and Al Collins) designed to demonstrate the use of goal hierarchies during reading (e.g., Experiment 1 discussed later). After a series of eight experiments, we could find evidence for the use of local goals but no evidence at all for the use of higher order goals. It was only much later, after several years and a number of other results (e.g., McKoon & Ratcliff, 1986), that we finally came to adopt the minimalist position.

The minimalist position contrasts with the framework that underlies most previous and current psychological investigations of inference processing during reading. Modern investigators began with the studies of Bransford and Franks and their colleagues, who adopted a strong constructionist approach to text processing (Bransford, Barclay, & Franks, 1972; Bransford & Franks, 1971; Johnson, Bransford, & Solomon, 1973). They interpreted their experimental results as demonstrating that encoding processes constructed inferences that were necessary to represent the situation described by a text. For example, a complete description of the sentence "Three turtles rested on a floating log, and a fish swam beneath them" would include the inference that the fish swam under the log. From the constructionist framework, this inference should be automatically encoded. From the minimalist position as proposed in this article, the inference would not be automatically encoded because it is not necessary to achieve local coherence, nor is the information that the fish swam under the log general knowledge.

Following Bransford et al.'s (1972) early work, constructionist hypotheses were advocated and tested by Richard Anderson and his colleagues (R. C. Anderson & Ortony, 1975; R. C. Anderson et al., 1976) and currently are embodied in some mental models approaches to text processing (Black & Bower, 1980; Glenberg, Meyer, & Lindem, 1987; Johnson-Laird, 1980; Mandler & Johnson, 1977; Morrow, Greenspan, & Bower, 1987; Rumelhart, 1977; Stein & Glenn, 1979; Trabasso & van den Broek, 1985; van Dijk & Kintsch, 1983). These models propose that the automatically encoded, mental representation of a text is a model of the situation described by the text. The representation is supposed to contain many nonminimal inferences, including elaborations on explicitly stated pieces of information and global connections among propositions. These constructionist models stand in direct opposition to the minimalist approach.

In this article, support for the minimalist position is provided in three ways. The first section of the article demonstrates a contrast between inferences that are constructed for local coherence and inferences that might be constructed to combine more global elements of a text. Several constructionist theories of text processing propose that global inferences are automatically constructed to connect pieces of information that are widely separated in a text; global inferences provide the overall structure of the text, such as the framework of a typical fairy tale or the causes of characters' actions. For local inferences, a review of recent research shows that several kinds are encoded during reading, as would be expected from a minimalist theory. In contrast, the results of Experiments 1 through 4 show that causal global inferences are not automatically encoded, in contradiction to some global theories.

A second body of research that supports the minimalist position is research that has examined elaborative inferences. These inferences represent information that is not required for local coherence. For example, semantic inferences might add contextually appropriate features of meaning to the representation of a concept, instrumental inferences might add the typical instrument for a verb (e.g., *spoon* for *stirring coffee*), and predictive inferences might add information about "what should happen next" in a story. A review of previous studies shows that, for instrumental and predictive inferences, the data contradict the constructionist hypothesis and support the minimalist hypothesis. For inferences about the contextually appropriate meanings of words, the data are consistent with both hypotheses.

Finally, several studies that examined the use of lifelike situation models during reading are considered. It has been proposed that a situation model represents textual information in a way that corresponds to a "real-life" situation (cf. Glenberg et al., 1987). For example, for a character described in a text as moving from one room to another, the situation model would automatically keep track of the character, associating the character first with the objects in one room, then the next room, and so on as the character moved (Morrow, Bower, & Greenspan, 1989). In the third section of this article, studies designed to demonstrate the automatic encoding of lifelike situation models are shown to have alternative interpretations, and a new experiment demonstrates the plausibility of one such interpretation. The alternative interpretations are consistent with the minimalist view, and no elaborated situation model is required.

The remarkable conclusion to be drawn from both the new experiments and the review of previous experiments is that the widely accepted constructionist view of text processing has almost no unassailable empirical support (see also Alba & Hasher, 1983). The constructionist view has been discussed and tested for the past 20 years. Yet, it is difficult to point to a single, unequivocal piece of evidence in favor of the automatic generation of constructionist inferences. In the General Discussion section, we suggest that future research should investigate a variety of kinds of inferences, aiming toward a deep understanding of the processing and informational bases of each kind. We suggest that such investigation will lead to a gradual expansion of the kinds of inferences identified as minimal: The immediately available information in short-term memory may be more complexly structured than originally supposed, and the immediately available information from general knowledge may be more varied than we now believe. It is the goal of the minimalist hypothesis to motivate this expansion.

It should be stressed that the minimalist and constructionist positions disagree on the question of what inferences are encoded automatically, as the basis for more strategic inferences or when readers do not have special goals and strategies, and that it is these automatic inferences that are the topic of this article. All of the inferences that might be (and often are) strategically generated as the result of special goals adopted by motivated readers are critically important to language understanding, problem solving, and learning. The minimalist position separates these inferences from minimal inferences, and so they are outside the scope of this article. However, at some point the connection must be made between the mental representations provided by minimal inferences and the processes that operate on them to form strategic inferences, and the issue must be addressed of how minimal inferences support other kinds of inferences. These problems are no less important than those described in this article.

Local Versus Global Coherence

The minimalist hypothesis makes an important distinction between the inferences that are required to establish local coherence and those that might connect more globally separated pieces of information. This distinction is not one that would be made from a constructionist viewpoint; a constructed representation of the situation described by a text would not necessarily include aspects of the situation that were mentioned in close proximity, and it would not necessarily exclude aspects that were more widely separated. However, in support of the minimalist position, the distinction is clearly apparent in the results of empirical studies. On the one hand, there is a large body of evidence favoring the hypothesis that local inferences are automatically generated. On the other hand, there is little evidence for the automatic generation of global inferences during reading, and Experiments 1 through 4 provide explicit evidence that one kind of global inference, causal inferences, is not generated.

Local Coherence

A major claim of the minimalist view is that inferences are constructed during reading to the extent that the information on which they depend is readily available. If the required information is not readily available, then an inference will not be constructed (unless the text is not locally coherent). An obvious potential source of readily available information is the information in short-term memory, and so it is hypothesized that inferences based on this information are automatically constructed. To support the minimalist position, it must be shown both that the supporting information is readily available and that the supported inference is encoded.

For the processing of text through short-term memory, we follow the model proposed by Kintsch and van Dijk (1978), although for other purposes we would update this model to the more complex representations of discourse models (cf. Greene, McKoon, & Ratcliff, 1992; Grosz, Joshi, & Weinstein, 1983; Sidner, 1983a, 1983b; Ward, Sproat, & McKoon, 1991; Webber, 1983). In Kintsch and van Dijk's model, the information in short-term memory during reading is assumed to be made up of explicitly stated words of the text plus the propositions that are being formed from them. The amount of information in short-term memory at any point in reading a text is loosely defined to be several clauses or sentences, depending on their length (cf. Daneman & Carpenter, 1980). The relevant issue for the current discussion is not an exact specification of the amount of information in short-term memory at any point in processing, but rather the contrast between information that can be described as being locally available and global information. When local inferences are examined empirically, they involve pieces of explicitly stated information that are close together in a text. When global inferences are examined empirically, they involve pieces of information that are so widely separated in the text that it is clear they could not be in short-term memory at the same time (without retrieval from long-term memory).

In the Kintsch and van Dijk (1978) model, the clauses in short-term memory are converted into semantic propositions. These propositions are connected together through overlap of their arguments, and they are ordered with respect to the most salient or topical proposition. Sentences and clauses do not usually provide an explicit representation of their underlying propositions and the connections among them; this information must often be inferred. For example, in the sentence "The mausoleum that enshrined the czar overlooked the square" the propositions are (roughly and informally) *mausoleum enshrined czar* and *mausoleum overlooked square,* where the two propositions refer to the same mausoleum. To form the appropriate locally coherent structure for the sentence, it must be encoded that the mausoleum both overlooked the square and enshrined the czar. In Kintsch and van Dijk's model, the processes that construct propositions are assumed to recognize that different occurrences of the same argument are in fact the same however the argument might be referenced (e.g., by a noun or an anaphor). Thus, the model assumes the encoding of the basic inferences necessary to form propositions through argument overlap. The minimalist view incorporates these inferences because they are based on the easily available information of short-term memory.

Empirical evidence confirms the assumption that inferences necessary to establish argument overlap are encoded. The encoding of inferences that establish propositional semantic units is well documented. Recall of a text depends on the number of propositions in the text (Kintsch & Keenan, 1973), and propositional units tend to be recalled as a whole (Kintsch & Glass, 1974). However, recall studies do not provide completely convincing evidence about encoded structures; with unlimited time in free recall, subjects may edit their responses to make them seem grammatical (i.e., by deleting incomplete propositions). Other evidence about propositional structures comes from priming studies with recognition memory. Ratcliff and McKoon (1978; see also McKoon & Ratcliff, 1980b; Ratcliff & McKoon, 1981b) gave subjects short lists of sentences to study. After each study list, subjects were given a recognition test list made up of single words from the sentences and unrelated distracter words. A subject's task was to decide as quickly as possible if each word in the test list had or had not appeared in a studied sentence. If a target test word from one of the sentences was immediately preceded in the test list by another word from the same sentence, then response time for the target was speeded. This priming effect was significantly greater if the two words from the sentence were from the same proposition than if they were from different propositions. For example, for the mausoleum sentence, response time for the target SQUARE was faster when *square* was primed by *mausoleum* from the same proposition than when it was primed by *czar* from the other proposition. These propositional priming effects have been shown to be due to automatic retrieval processes (Ratcliff & McKoon, 1981a), indicating that the structures reflected by priming were encoded during reading.

Evidence that inferences to establish propositional units are encoded during reading supports the minimalist position only if it can also be shown that the information on which the inferences are based is easily available. Studies that indicate immediate availability are provided in recent work by Swinney and his colleagues (cf. Nicol & Swinney, 1989; Swinney & Osterhout, 1990), who used a cross-modal on-line lexical decision task. They used sentences like "The policeman saw the boy that the crowd at the party accused of the crime." In this sentence, *boy* should be encoded as the person who was accused (in the proposition *crowd accused boy*), and so *boy* should be quickly available after the word *accused.* To test this, sentences were presented auditorily, and at various points during the sentences, lexical decision test items were displayed visually. The lexical decision test items were strong associates of critical words in the sentences. The reasoning was that there should be facilitation in response time for an associate at any point where its related critical word was being used in comprehension. For example, the lexical decision for an associate of *boy* should be facilitated after the word *accused* because *boy* is the object of *accused.* The data showed this result and also that the associate was not facilitated after the word *party,* a point in the sentence where *boy* would not be used in building the underlying structure of the sentence. Similar evidence of immediate availability has been reported by Tanenhaus, Carlson, and Seidenberg (1985) and by Garnsey, Tanenhaus, and Chapman (1989). This evidence is all consistent with the idea that the information necessary to make connections among propositions is quickly available. The total combination of evidence—that inferences about propositional connections are encoded (Ratcliff & McKoon, 1978) and that the information on which they depend is quickly available (Nicol & Swinney, 1989)—exactly fits the minimalist hypothesis.

A second kind of inference that is often needed to establish argument overlap is the connection between an anaphor and its referent. If a text mentions some pronoun and predicates information about the pronoun, then the information about the pronoun should be connected to a referent of the pronoun and to other information given by the text about that referent. The processing of coreference has been extensively studied. For example, Corbett and Chang (1983; also Chang, 1980; Clark & Sengul, 1979; Ehrlich & Rayner, 1983) used sentences like "Rachel tried to catch Sally, but she was not able to do it," with the possible referents of *she* presented for recognition test at the end of the sentence. They found that responses to the intended referent were faster than responses to the unintended referent (but see Gernsbacher, 1989; Greene, McKoon, & Ratcliff, 1992). Nicol (1988, cited in Nicol & Swinney, 1989) has demonstrated the availability of potential referents of pronouns more immediately than at the end of sentences. She used a cross-modal on-line lexical decision task (as presented earlier), and sentences like "The boxer told the skier that the doctor for the team would blame him for the recent injury." When test words were presented immediately after the pronoun *him,* there was facilitation of response times for associates of the potential referents of the pronoun (*boxer* and *skier*). However, there was no facilitation for an associate of the noun that could not be a referent (*doctor*). This pattern of data is consistent with information about potential referents being quickly available, and so the result is consistent with the minimalist hypothesis.

A more stringent test of the minimalist position would be a combination of studies that showed both the encoding of appropriate connections between referent and anaphor and the immediate availability of the information that supports the connections. Such studies have not been done for pronouns, but they have been done for nominal anaphors (Dell, McKoon, & Ratcliff, 1983; McKoon & Ratcliff, 1980a). These experiments used short texts that, in the first sentence, mentioned a character such as a burglar. "A burglar surveyed the garage set back from the street. Several milk bottles were piled at the curb. The banker and her husband were on vacation. The criminal/A cat slipped away from the streetlamp." In the last sentence, either the character introduced in the first sentence was referenced again with a category label (*the criminal*), or a new character (*a cat*) was introduced, with no mention of the character from the first sentence. When the last sentence referred to the burglar as *the criminal,* information about the burglar should have been directly connected from the first sentence to the last sentence. McKoon and Ratcliff (1980a) showed that these connections were encoded using recognition priming. Subjects were given study lists of texts to read. After each study list, they were given single words for recognition. Among the test words was a noun from the last sentence, and it was immediately preceded in the test list by the character from the first sentence (e.g., *streetlamp* immediately preceded by *burglar*). When the noun and the character were directly connected together in the text by the anaphor (*The criminal*), response times on the noun were speeded relative to when the noun and the character were not directly connected (when it was the cat that slipped away from the streetlamp). This result shows the encoding of connections based on anaphoric inferences.

Results indicating the immediate availability of information supporting the connections were obtained by Dell et al. (1983). They used a word-by-word reading procedure in which each word of a text was displayed for 250 ms, and recognition test words could be presented after any word of the text. One test point was immediately after the first noun of the last sentence (the anaphor *criminal* or the word *cat*). At this test point, response times to the antecedent (*burglar*) and to another word from the same proposition as the antecedent (*garage*) were both facilitated in the criminal version of the last sentence relative to the cat version, consistent with immediate availability of the referent for the anaphor. Corbett (1984) also found results that indicate the immediate availability of potential referents for anaphors using a different paradigm. He found that reading times for anaphors like *wooden toy* were faster when there was only one possible referent in the text (*wooden block*) than when there was also a nonreferent from the same general category (*rubber ball*). Thus, taken in combination, these studies support the minimalist hypothesis by showing that the information necessary to establish anaphoric connections is available immediately during reading.

In the van Dijk and Kintsch (1983) processing model, the propositional connections established by repetitions of concepts and anaphoric relations are the only means of establishing local coherence. However, as Kintsch and van Dijk point out, propositional connections are not sufficient to guarantee coherence. Keenan, Baillet, and Brown (1984) made this point with the sentence pair "Tom Jones plans to go to the dentist. A plane flew over Tom Jones." According to the minimalist position, inferences will be encoded if they are required for local coherence. The problem is to define exactly what constitutes local coherence. No formal definition is available, although researchers have made several suggestions. Lack of a formal definition does not mean that local coherence cannot be investigated empirically. Other concepts in psycholinguistics that lack formal definitions (such as *proposition*) have been used to excellent advantage (cf. Kintsch & Keenan, 1973; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Kintsch & van Dijk, 1978), and empirical investigation should lead to more formal descriptions and definitions of local coherence. For present purposes, we assume that a set of two or three sentences is locally coherent if it makes sense on its own or in combination with easily available general knowledge. It is not locally coherent if information from elsewhere in the discourse is required.

Suggestions for the kinds of inferences that might be involved in local coherence include bridging inferences and causal inferences. Haviland and Clark (1974) outlined several kinds of bridging inferences, and Keenan and Kintsch (1974; also McKoon & Keenan, 1974) provided data to indicate that bridging connections are encoded into the memory representation of a text. An example of a text used by Keenan and Kintsch is "Police are hunting a man in hiding. The wife of Bob Birch disclosed illegal business practices in an interview on Sunday." For this text, a bridging inference is required to provide the relation between Bob Birch and the man in hiding. Keenan and Kintsch found evidence that this inference is encoded during comprehension. They used a verification test (given 15 min after the text was read). Response times for the statement "Bob Birch is the man who is hiding" were just as fast for the text that required the bridging inference as for another version of the text that made the inference explicit. From this result, Keenan and Kintsch argued that this kind of bridging information was encoded during reading. Whether the result is fully consistent with the minimalist position is not clear. The information that Bob Birch is the man who is in hiding is not known before reading the text, and so it would not be quickly and easily available. Therefore, the minimalist prediction would be that it was constructed by a relatively slow inference process; this prediction has not been tested.

Another potential contributor to local coherence is causality; propositions that are in short-term memory at the same time have been said to be connected by their causal relations. One way to demonstrate the importance of causal relations would be to show that causally relevant propositions are preferentially maintained in short-term memory during reading. Fletcher, Hummel, and Marsolek (1990) found evidence for such maintenance, although it could be argued that, with their materials, causally relevant propositions were maintained in short-term memory by virtue of (anaphoric) repetitions of their content rather than by virtue of their causality.

Other demonstrations of the effects of causal relations have used pairs of sentences that were designed to vary in their causal relatedness. Keenan et al. (1984; see also Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990; Myers, Shinjo, & Duffy, 1987) found that the reading time for the second sentence of a pair was slowed as the causal relatedness of the pair was decreased. There are two possible interpretations of this result: One is that reading time was slowed by the process of constructing (or attempting to construct) a causal chain to relate the two sentences—less related sentences require the construction of a longer chain. The other interpretation is that reading time slowed because of difficulty in finding an already existing causal chain in long-term memory. By this interpretation, closely related sentences are causally connected through a relation provided by long-term memory. The causal chain that connects two closely related sentences may be long or short, but it will be quickly processed because it is already available and does not have to be constructed. Less closely related sentences would represent a mixture of processes, some connected by difficult-to-access relations in long-term memory, some connected by newly constructed relations, and some perhaps left without any causal connection.

Given these different interpretations, it is not clear whether the causal connections investigated in these studies were encoded automatically. From the minimalist point of view, the causal relations encoded automatically during reading should be those that are quickly available from long-term memory; those that are not available from long-term memory but are required to establish local coherence should also be encoded. This claim has not been tested empirically. One problem is to define what causal inferences are necessary for coherence; we return to discussion of this problem after considering research on global inferences.

Leaving aside the uncertain situation with causal relations, the minimalist hypothesis is well supported with respect to local coherence: Current data are consistent with the claims that inferences based on quickly available information are encoded during reading. The minimalist position would be contradicted if it could be shown that some inference was encoded even though it was neither quickly available nor necessary for local coherence. The minimalist position would also be contradicted if it could be shown that there were kinds of quickly available information that did not support inferences. However, there is no such evidence to contradict the minimalist claims. In the next section, we show that the situation for global inferences is much different than that for local inferences. Although the local inferences for propositional structures posited by the minimalist view are relatively easy to demonstrate empirically, there is no evidence that global inferences for global structures are automatically generated during comprehension.

Global Inferences

Many researchers have proposed that global inferences connect widely separated pieces of textual information and that they do so automatically as a necessary part of comprehension. Sometimes these inferences are analyzed as the linking elements of a story "grammar" so that initiating settings, characters, goals, and events are linked to their consequent events and outcomes (Mandler, 1978; Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1979; Thorndyke, 1977). More often, global inferences are the links that connect explicit pieces of information into an overall causal chain or network (Black & Bower, 1980; Graesser, 1981; Graesser, Robertson, & Anderson, 1981; Omanson, 1982a, 1982b; Trabasso & van den Broek, 1985; Trabasso & Sperry, 1985). From the minimalist point of view, these inferences should not be automatically constructed during reading. They are usually not required to establish local coherence, and they are usually not supported by well-known information. Only if a text is locally incoherent at some point should global information be recruited to establish local coherence. Of course, readers will often construct global inferences when such inferences are required by the readers' goals. Minimalist inferences will be constructed in the absence of special goals or strategies and to provide the bases for goal-driven inferences.

Experiments 1 through 4 examined whether global causal inferences are generated automatically during comprehension. Because the experiments directly challenge the hypothesis that global inferences are encoded automatically, it is necessary to explain clearly what kinds of inferences are both causal and global. As an illustration, we use the method of analysis of causal relations developed by Trabasso and his colleagues (cf. Trabasso & van den Broek, 1985).

Table 1 shows a short story and its analysis, adapted from an article by Suh and Trabasso (1988). The meaning of each sentence in the story is identified as setting, initiating event, goal, action, outcome, or reaction. These elements make up the definition of an *episode*. For an episode to occur, there must be a setting in which it occurs, one or more initiating events in the setting, and reactions to the events. If the reactions lead to a goal, then one or more actions will result, and they in turn will have outcomes. This episode structure is recursive in that outcomes may provide the initiating events for further reactions, goals, and outcomes. The definition of the episode structure requires that each goal be linked directly to its initiating event or events and each outcome be linked directly to the goal it fulfills (or fails to fulfill). It is assumed that these direct links must be encoded during reading. If the links are not explicitly stated, then they will be inferred. If the necessary pieces of information to create the links are not locally available, then they will be retrieved from memory. The links between initiating events and goals, and between goals and outcomes, that are assumed for the story in Table 1 are shown at the bottom of the table. The mother's birthday is the initiating event for the goal of wanting to buy a present, and the outcomes of this goal are that everything was too expensive and no present was bought. These outcomes plus the original initiating event, the birthday, provide the initiating events for the second goal, knitting a sweater. For this second goal, an inference is required, namely, that the sweater was to be the mother's birthday present. This is labeled a *global inference* if it is the case that the initiating event, the mother's birthday, is no longer available in working memory when the second goal is read. The specific analysis for the story in Table 1 is from Trabasso and van den Broek (1985), but other causal analyses (e.g., Black & Bower, 1980; Graesser, 1981; Mandler & Johnson, 1977; Omanson, 1982a, 1982b; Rumelhart, 1975; Stein & Glenn, 1979; Thorndyke, 1977) would also assume that the inferred link between the birthday and knitting the sweater was encoded during reading into the mental representation of the story.

A number of empirical results have been obtained that are consistent with causal analyses of stories. The largest body of

Table 1
A Short Story From Suh and Trabasso (1988)

Setting: *Once there was a girl named Betty.*
Initiating Event 1: *One day, Betty found that her mother's birthday was coming soon.*
Goal 1: *Betty really wanted to give her mother a present.*
Action: *Betty went to the department store.*
Outcome 1: *Betty found that everything was too expensive.*
Outcome 2: *Betty could not buy anything for her mother.*
Reaction: *Betty felt sorry.*
Initiating Event 2: *Several days later, Betty saw her friend knitting.*
Setting: *Betty was good at knitting too.*
Goal 2: *Betty decided to knit a sweater.*
(*Story continues*)
Goal 1 is linked directly to its Initiating Event 1.
Outcomes 1 and 2 are linked directly to their Goal 1.
Goal 2 is linked directly to its initiating events, which are Initiating Event 1 and Outcomes 1 and 2.

data comes from recall studies. The probability of recalling any particular fact can be predicted from its position in the causal network representation of its story (cf. Black & Bower, 1980; Omanson, 1982a; van den Broek, 1988). Causal information that is on a direct causal chain from the beginning of a story to the end is more likely to be recalled than information that is not on the chain (Trabasso & van den Broek, 1985; van den Broek & Trabasso, 1986). Also, the probability of recalling a piece of information increases with the number of causal connections it has to other pieces of information (Trabasso & Sperry, 1985).

These recall findings have often been cited as support for the hypothesis that global causal inferences are encoded during reading. However, recall does not necessarily measure encoding. It may be that recall sometimes gives an accurate measure of encoded information, but it may also measure the results of the retrieval and editing processes that operate on encoded information, and these processes may give nonrandom distortions of the encoded information. For recall of stories, it is easy to see that subjects might edit the facts of their encoded representations into causally connected structures, eliminating facts that they remembered but decided not to write down and working extra hard to remember facts that would turn an otherwise unrelated list of sentences into a coherent story. Thus, the causal structures found in recall protocols may be a reflection of editing processes and not an accurate reflection of the representation in memory that was formed by encoding processes.

This point is reinforced by empirical demonstrations of the roles of retrieval and editing processes. For example, Alba, Alexander, Hasher, and Caniglia (1981) showed that subjects could recognize statements from stories for which they knew the topic as well as they could recognize statements from stories for which they did not know the topic, even though recall was much worse when they did not know the topic. Another clear example of the operation of retrieval processes is provided in a study by Singer (1978). He showed that the effectiveness of a cue for recall was determined by backward associations at the time of recall from the cue back to the text to be recalled, not forward associations inferred when the text was read. Other results by Corbett and Dosher (1978) and Baillet and Keenan (1986) also demonstrate that recall experiments do not provide convincing evidence that inferences are generated during reading.

The processes that can be involved in recall, including editing and inference generation, are important processes to study, but they are not the focus of this article. Our aim is to separate out and focus on the inferences that are automatically included in a text representation at encoding. In this way, a clearer demarcation can be drawn between processes that occur at encoding and those that can occur at retrieval.

In Experiments 1 through 4, we use experimental procedures other than recall to compare causal global inferences to inferences based on locally available information. From the minimalist hypothesis, we expected that global inferences would not be automatically encoded during reading. This finding is also predicted by results from experiments by Glanzer, Fischer, and Dorfman (1984). They interrupted subjects' reading in the middle of a text and gave them an unrelated task to perform. When the subjects resumed reading the text, the best aid to comprehension was not global information about the topic of the text, but local information from the context immediately preceding the interruption.

Empirical Tests for Global Causal Inferences

The basic hypothesis that runs through all of Experiments 1 through 4 is that, barring special strategies by readers, causal global inferences are not constructed if a text is locally coherent. Only when a text is not locally coherent will global information be brought in to aid comprehension. Of course, readers can and often do adopt special strategies, either during reading or recall, to involve global information in local processing. However, in the typical laboratory experiment without special instructions, such strategies do not appear to be used during reading.

The hypothesis that global inferences are not automatically constructed for locally coherent texts is suggested by consideration of simple examples. Suppose a story relates that, when a killer's rifle won't work properly, he reaches for his hand grenades. This sequence of events makes sense without global knowledge of the killer's goal, to assassinate a president. On the other hand, if a text is not locally coherent, then global information should be used. When a character in a story decides to buy fruit and yogurt as a result of finding her bicycle broken, a reader needs the global information that she is trying to lose weight to make sense of the scenario.

Experiment 1 contrasted the availability of local and global information during reading of short texts. Causal global inferences were identified using the definitions given by Trabasso (Suh & Trabasso, 1988; Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985) and described earlier. All of the texts were locally coherent, and results indicated that local information is available during comprehension. The texts did not require global causal information for coherence at the local level, and results indicated that it was not used. Experiment 2 extended these results with texts of two types. One type was coherent at the local level, but local information contradicted global information. The data showed no effects of this contradiction. The second type of text was not coherent locally, although it could be made coherent through global information. In this case, the data showed that global information does become available for use at the local level.

Experiments 3 and 4 used long, naturalistic stories to investigate the representations of inferences in memory. The cause of some specific event in a story was separated from the event by several paragraphs. The empirical question was whether global inferences would connect the event to its cause in the memory representation of the story. The data indicated that this does not happen. Thus, over all four experiments, there is no evidence that global causal inferences are constructed during reading.

Experiment 1

This experiment was designed to assess the availability of local and global information at the end of reading short texts. Each text had two paragraphs, an introduction paragraph and a continuation paragraph, as shown by example in Table 2. In the introduction, a general goal (e.g., killing the president) and a goal subordinate to the general goal (using a rifle) are described. For the continuation paragraph, there were three different versions: Control, Try Again, and Substitution. In the Control continuation, both goals are achieved (the president is shot), and a new goal is introduced. In the Try Again continuation, a problem arises in achieving the subordinate goal and the character tries this goal again (using the rifle in a different way). In the Substitution continuation, a problem also arises with the subordinate goal, but instead of trying again, the character replaces it with a new subordinate goal (hand grenades). The new subordinate goal, like the old one, is designed to achieve the original general goal (killing the president).

Subjects read each text one sentence at a time, at a pace they controlled themselves. Availability of a goal was tested by presenting a recognition test word for the goal immediately after the final sentence of the text.

For the general goal in the texts, the minimalist and constructionist positions make different predictions. All of the continuations were written to be coherent in themselves; the general goal is not needed to comprehend any of them. Thus, according to the minimalist prediction, the general goal should not be used during comprehension of any of the continuations, and so the availability of the general goal should be equal across the different continuations. Responses to the general goal test word should not differ across the continuations in speed or accuracy. In contrast, according to a constructionist theory, responses to the general goal test words should be faster in the Try Again and Substitution continuations than in the Control continuation. This is because the character in the text is still trying to achieve the general goal at the end of both the Try Again and Substitution conditions but not at the end of the Control condition (where a new general goal has taken over).

For the subordinate goal, the minimalist and constructionist positions can make the same predictions. Locally, the original subordinate goal is necessary for comprehension only in the Try Again continuation; in neither the Control condition nor the Substitution continuations is the original subordinate goal still necessary to understand the character's actions. Thus, responses to the subordinate test word should be faster, more accurate, or both in the Try Again condition relative to the other two. For a constructionist theory, the character is still trying to achieve the original subordinate goal in the Try Again continuation, and so responses should be facilitated in this condition relative to the Control. In the Substitution continuation, there might or might not be facilitation, depending on whether the switch to a new subordinate goal eliminated all facilitation for the original subordinate.

Table 2
An Example of a Story From Experiment 1

Part of story	Story
Introduction	The crowd's cheers alerted the onlookers to the president's arrival.
	The assassin wanted to kill the president.
	He reached for his high-powered rifle.
	He lifted the gun to his shoulder to peer through its scope.
Control continuation	The assassin hit the president with the first shot from his rifle.
	Then he started to run toward the west.
	The searing sun blinded his eyes.
Try again continuation	The scope fell off as he lifted the rifle.
	He lay prone to draw a sight without the scope.
	The searing sun blinded his eyes.
Substitution continuation	The scope fell off as he lifted the rifle.
	So he reached for his hand grenades.
	The searing sun blinded his eyes.
General goal test word: *Kill*	
Subordinate goal test word: *Rifle*	

Note. The labels of the parts of the stories were not presented to the subjects.

Method

Materials. Each of the 30 experimental texts was made up of an introduction and three different continuations. The introduction introduced a general goal for the main character in the story (e.g., killing the president for the text in Table 2) and a subordinate goal that was a way of obtaining the general goal (e.g., using a rifle). The general goal was mentioned only once in the introduction and was not mentioned explicitly in the continuations. The subordinate goal was mentioned once in the introduction and again in the first sentence of each continuation. The introductions were always four sentences in length. The general goal and the subordinate goal were used as test words (e.g., *kill* and *rifle*).

In the Control continuation, the first sentence described successful fulfillment of the subordinate goal and so, by implication, the general goal. Then, the second sentence described a new general goal for the character. Examples of the original general goals in the introductions and new general goals in the Control continuations include going out for an evening's entertainment and then finding out where to buy furniture, cleaning house and then painting a barn, eating and then back scratching, getting a front-page story and then moving a printing press, investing money and then stopping at a dry cleaners, and holding a sale and then going to Europe.

The second continuation, the Try Again condition, described a problem with fulfilling the subordinate goal and presented a new method for fulfilling the same subordinate goal. Examples of the new and old methods include having a lecture from a doctor instead of from a social worker, going somewhere by train instead of by car, adopting a baby through a lawyer instead of an agency, borrowing money from a bank instead of a relative, asking a sister to do something instead of a friend, and going to cheerleader practice instead of going home. Note that in each case, the two methods of achieving the subordinate goal are coherent alternatives even though the general, superordinate goal for these examples is not given here. For example, having a lecture from a doctor instead of a social worker makes sense without knowing that the general goal is to obtain information about the world's population problems.

The third continuation, the Substitution condition, also described a problem with fulfilling the original subordinate goal and presented a new subordinate goal that would fulfill the original general goal. Examples include raking the lawn instead of trimming the hedge, a lecture about the world's food supply instead of about birth control, going to see fireworks instead of to the beach, giving money to a charity instead of to specific people, selling stocks instead of borrowing money, and going to a night club instead of to a movie. In each of these examples, the alternative makes sense without the general goal. For example, trimming the hedge can be understood as a substitute for raking the lawn without knowing the general goal of getting ready for a lawn party.

Each continuation was three sentences in length, and the final sentences of the three continuations were identical. The continuations were all locally coherent, as shown by the examples, in that they could make sense without knowledge of the general goal stated in the introduction.

In some of the Try Again and Control continuations, the second sentence contained words that might be semantically associated (preexperimentally) to one or the other of the test words. For example, the words *sight* and *scope* are associated with rifle. However, the number of items with such associations was about the same for the Try Again and Substitution continuations for both the general goal and subordinate goal test words. Thus, overall, associations between words in the texts and test words were equated across all conditions but the Control.

There were also 42 filler texts, each with one test word. Nine of the fillers were five sentences in length, 9 were six sentences in length, and 9 were eight sentences in length. For each length, six of the tests had positive test words and three had negative test words. The other 15 texts were seven sentences in length and had negative test words.

Procedure. The presentation of stimuli and collection of responses was controlled by a real-time computer system. Stimuli were displayed on a cathode-ray tube (CRT) screen, and responses were made by pressing keys on the CRT's keyboard.

The experiment began with a practice list of 8 texts, each one to three sentences in length. Then the 72 texts of the experiment proper were presented, eight fillers first and then the remaining texts in random order.

Presentation of each text began with an instruction displayed on the CRT screen asking the subject to press the space bar. When the space bar was pressed, there was a 200-ms pause, and then the first sentence of the text was displayed. The sentence remained on the screen until the subject pressed the space bar again; then the screen was cleared, there was a 500-ms pause, and then the next sentence of the text was displayed. Presentation of the sentences continued in this way until the final sentence of the text. After the final sentence was displayed and the space bar pressed, a row of asterisks appeared with a test word immediately below it. The subjects' instructions were to indicate whether the test word had appeared in the immediately preceding text, by pressing the "?/" key for a positive response and the *z* key for a negative response. The test word was erased from the screen immediately after the response. If the response was incorrect, the word ***ERROR*** was presented for 2,000 ms, and then the screen was cleared and there was a pause of 200 ms. If the response to the test word was correct, then there was a 200-ms pause. After the pause, the response time for the test word was displayed for 800 ms, then there was a 500-ms pause, and the instruction to press the space bar to begin the next text was displayed. Subjects were instructed to read the texts carefully and to respond as quickly and accurately as they could to the test words.

Design and subjects. With one group of 18 subjects, the test word for the experimental texts was always the general goal; for the other group of 18 subjects, it was always the subordinate goal. The experimental texts were presented with the Control, the Try Again, or the Substitution continuations. This variable was combined with three sets of subjects (6 per set in each group) and three sets of texts (10 per set) in a Latin square design. The subjects participated in the experiment for credit in an introductory psychology course.

Results

The mean reading time for the final sentence of each text and the mean response times and error rates for each subject and each test word were calculated; means of these means are displayed in Table 3. There were no specific predictions about final sentence reading times; they are included for completeness. Subjects who were tested with the subordinate goal test words read faster than subjects who were tested with the general goal test words.

According to the minimalist local coherence position, only local information and not the original general goal is necessary for comprehension of the continuations. Thus, response time and accuracy for the general goal test words should not vary across the different continuations, as is shown in the data in Table 3. For the subordinate goal test words, in the Control and Substitution continuations, a new goal was substituted so that the original subordinate need no longer be involved in comprehension at the end of the continuations; as a result, response times for the subordinate test word should be relatively slow and/or inaccurate. In the Try Again continuation, the original subordinate goal was still necessary for comprehension of the character's actions, so response times should be relatively fast and/or accurate. This is the pattern of data shown in Table 3.

For the test words expressing the original general goal, analysis of variance (ANOVA) showed no significant differences in response times, error rates, or reading times for the final sentences. For the test words expressing the original subordinate goal, the response times were significantly different across the continuations, $F(2, 34) = 5.5$, with subjects as the random variable, and $F(2, 58) = 3.4$, with test words as the random variable. The standard error of the response times was 11.6 ms. There were no significant differences in error rates or in the reading times of the final sentences (Fs < 1).

Discussion

The critical comparison between the minimalist local coherence hypothesis and the global constructionist hypothesis rests in their predictions for the general goal test word. According to the constructionist hypothesis, the character in the text is still trying to achieve the general goal at the end of the Try Again and Substitution continuations, and so responses to the goal test words should be facilitated in these conditions relative to the Control condition. According to the minimalist hypothesis, responses for the general goal test words should not differ across the three conditions because all the continuations are locally coherent and none require the general goal for local

Table 3
Results From Experiment 1: Mean Response Times (in Milliseconds), and Error Rates for Test Words and Mean Reading Times (in Milliseconds) for Final Sentences

	General goal tested			Subordinate goal tested		
	Test words			Test words		
Type of continuation	RT	% error	Reading times	RT	% error	Reading times
Control	717	11	1,551	638	5	1,399
Try again	717	12	1,588	594	8	1,399
Substitution	718	12	1,585	644	7	1,337

Note. RT = response time.

comprehension. The data support the minimalist view because there are no differences across the conditions. The data for the subordinate goal test words do show differences across conditions, indicating that the experiment did not lack power.

Experiment 2

Experiment 2 was devised to provide additional tests for global inferences. The procedure was the same as in Experiment 1: Subjects read short texts sentence by sentence, and recognition test words were presented after the final sentence. Examples of the texts are shown in Table 4. There were two kinds of texts, Globally Inconsistent and Locally Inconsistent, each with an Introduction plus a Control continuation and a Problem continuation.

The first text is labeled *Globally Inconsistent.* This reflects the fact that, in the Problem continuation, watching videotapes is not consistent with the stated goal of working out an injured arm. The text provides a test for global inferences because the inconsistency should amplify the use of global information at the local level, and so responses to the test word *workout* should be facilitated relative to the Control condition. However, the Problem continuation, like the Control continuation, is locally coherent; neither requires use of the general goal for comprehension. If only local information is used in comprehension, then there should be no facilitation of *workout* in the Problem condition relative to the Control condition.

The second type of text is labeled *Locally Inconsistent* because replacing a broken bicycle with grapefruit and yogurt does not make sense on the local level. However, it does make sense in the global context of trying to lose weight. For this text, both the global inference and the minimalist positions agree: The global goal information about losing weight should be recruited during local processing, and responses to the goal test word (*weight*) should be facilitated in the Problem condition relative to the Control condition.

Method

Materials. Each of the experimental texts used in the experiment was made up of an introduction and two different continuations. The introductions, always four sentences in length, described some goal for the main character of the story (a workout in the first example in Table 4). This goal was mentioned explicitly only once in the introduction and not mentioned explicitly in either continuation. One word expressing the goal (e.g., workout) was used as the test word for the text. In the first continuation, the Control condition, the goal was fulfilled and a new goal described (the Control versions were similar to the Control versions used in Experiment 1). In the Problem continuation, some problem that prevented attainment of the original goal was described, and then a new goal was substituted. The final sentences of the two continuations were always the same, and all continuations were three sentences long.

There were two sets of experimental texts (20 in each set) that differed in the relation between the substitute goal in the Problem continuation and the original goal. In the Globally Inconsistent set of texts, the new goal was inconsistent with the original goal; some examples of new and original goals include fixing a lock in the attic instead of preparing the grounds for a lawn party, going to a restaurant instead of on a picnic, buying a conservative gown instead of buying something to look unusual, donating money instead of finding a cure for loneliness, watering the chickens instead of cleaning the house, buying a heated swimming pool instead of saving on electric bills, flying to Las Vegas instead of investing wisely, and serving take-out hamburgers instead of a sumptuous feast. In each case, the substituted goal cannot lead to achievement of the original goal—there is no way that take-out hamburgers can provide a sumptuous feast, and presenting the two goals in conjunction, as is done here, makes the inconsistency clearly apparent. What makes the inconsistency not obvious to readers of the texts is that the two goals are not simultaneously available. The continuation becomes locally coherent because there is a plausible relation between the problem and the substitute goal (e.g., take-out hamburgers are a plausible alternative when someone forgets to buy steak).

In the Locally Inconsistent set, the substitute goal was consistent with the original goal (as dieting is another way to lose weight in Table 4), but the relation between the problem (a broken bike) and the substitute (buying grapefruit and yogurt) could not easily be determined at a local level. Some examples of problems and the actions that resulted from them include going to McDonald's after finding a stopped clock, looking for a scarf when the power goes out, substituting a quilt for a clock, calling customers when the vegetables are overcooked, and looking in the cupboards when the car won't start. In these examples, it is not clear why the action results from the problem because the general goal is not given. For example, the general goal connecting the quilt to the clock was the search for something decorative to place above a fireplace mantel.

In addition to the test word that described the goal introduced in the introduction, each text also had two other test words, one for each continuation. About half of these were words that appeared in the continuation (positive test items), and about half were words that did not appear in any text at all (negative test items).

There were also 40 filler texts. Each was seven sentences in length, and each had two test words. Of all the filler test words, 30 were posi-

Table 4
Examples of Stories Used in Experiment 2

Part of story	Story
	Globally inconsistent
Introduction	Curtis spied a tennis court in the park. His arm was healing from an injury and needed a workout before the big match. So he needed an opponent. Curtis waved to a friend to join him.
Control continuation	The friend came over and was an exhausting opponent. Curtis decided to go borrow some change for a drink. Curtis ran happily along the path.
Problem continuation	Curtis' friend did not want to be Curtis's opponent. So Curtis decided to go home and study videotapes of his serve instead. Curtis ran happily along the path.
Goal test word: *Workout*	
	Locally inconsistent
Introduction	Diane wanted to lose some weight. She thought she should lose at least 20 pounds. Diane thought cycling might help her lose some weight. She went to the garage to find her bike.
Control continuation	Diane peddled 5 miles each day for 3 months and became very slim. She decided to go back to school to complete her degree. It took several years, but Diane finally reached her goal.
Problem continuation	Diane's bike was broken and she couldn't afford a new one. So she went to the grocery store to buy grapefruit and yogurt. It took several years, but Diane finally reached her goal.
Goal test word: *Weight*	

tive words from their texts and 50 did not appear in any text. The positive words were always chosen from the latter halves of their texts (because the goal test words from the experimental texts were always from the beginning of their texts).

Procedure. The presentation of stimuli and collection of responses were controlled by a real-time computer system. Stimuli were displayed on a CRT screen, and responses were indicated using keys on the CRT's keyboard.

The experiment began with a practice list of 30 texts, each one or two sentences in length. Then the 80 texts of the experiment proper were presented in the same manner as the practice texts.

Presentation of each text began with an instruction displayed on the CRT screen asking the subject to press the space bar. When the space bar was pressed, there was a 500-ms pause and then the first sentence of the text was displayed. The sentence remained on the screen until the subject pressed the space bar again; then there was a 50-ms pause, the screen was cleared, there was another 50-ms pause, and then the next sentence of the text was displayed. Presentation of the sentences continued in this way until the final sentence of the text. After the final sentence was displayed and the space bar pressed, a row of addition signs appeared with a test word immediately below it. The subjects' instructions were to indicate whether the test word had appeared in the immediately preceding text by pressing the "?/" key for a positive response and the *z* key for a negative response. If the response was incorrect, the letters of the word *ERROR!!* were presented one at a time for 600 ms each, and then the screen was cleared and a row of addition signs and a second test word were presented. If the response to the first test word was correct, then the test word was erased from the screen, there was a 100-ms pause, and then the row of addition signs with the second test word appeared. If the response to the second test word was correct, the instruction to press the space bar to begin the next text was displayed. If the response was incorrect, the error message was presented before the instruction to begin the next text. The order of presentation of the texts was randomly chosen, a different randomization for each second subject. For the experimental texts, the first test word was always the word expressing the goal mentioned in the introduction. For the filler texts, the correct response for the first test word was always negative. Subjects were instructed to read the texts carefully and to respond as quickly and accurately as they could to the test words.

Design and subjects. For one group of 50 subjects, the experimental texts were the Globally Inconsistent set, and for a second group of 50 subjects, they were the Locally Inconsistent set. Each of the sets was divided into two subsets. The subsets were combined in a Latin square design with two sets of subjects (25 subjects per set) and the two continuations, Problem and Control.

Results

Mean response times and error rates for the test words were calculated for each subject and each test word, and mean reading times were calculated for each sentence of each text. Means of these means are shown in Table 5.

From both the minimalist and global inference points of view, the Problem continuations of the Locally Inconsistent texts should require the use of global information. The original goal is needed for the continuations to be understood. Thus, responses to the general goal test word should be faster in the Problem Condition than the Control condition, which is what

Table 5
Experiment 2: Reading Times for Sentences and Correct Response Times (in Milliseconds) and Error Rates for Test Words

Data	Control continuation	Problem continuation
	Locally inconsistent texts	
Reading time		
Sentences 1–4	2,166	2,184
Sentence 5	2,000	1,973
Sentence 6	2,371	2,120
Sentence 7	1,524	1,567
Goal test word	1,086[a]	1,030[b]
	Globally inconsistent texts	
Reading time		
Sentences 1–4	2,078	2,144
Sentence 5	1,951	2,019
Sentence 6	2,454	2,345
Sentence 7	1,609	1,681
Goal test word	1,137[c]	1,164[d]
	Filler test word	
Positive test word	889[a]	
Negative test word	1,004[b]	

[a] Percentage error was 6 for this entry. [b] Percentage error was 4 for this entry. [c] Percentage error was 5 for this entry. [d] Percentage error was 8 for this entry.

the data show. For the Globally Inconsistent texts, the two points of view make different predictions; according to the local coherence position, there is no local problem with comprehension, and so there should be no significant difference between mean response times for the goal word in the Problem and Control continuations. According to the global coherence position, the general goal should still be involved in comprehension in the Problem Continuation, and so response times for the test word should be facilitated. As Table 5 shows, no significant facilitation was observed (the nonsignificant difference is in the wrong direction).

The results just described represent an interaction shown significant by ANOVA, $F(1, 49) = 4.68$, with subjects as the random variable, and $F(1, 38) = 7.32$, with test words as the random variable. There was also a significant main effect, that responses for the goal test words were slower for the Globally Inconsistent than the Locally Inconsistent texts, $F(1, 49) = 41.2$, with subjects as the random variable, and $F(1, 38) = 6.29$, with test words as the random variable. The standard error for the response times was 17 ms.

Post hoc tests showed the advantage for response times for the goal word with the Locally Inconsistent Problem texts to be significant, $F(1, 49) = 5.65$, with subjects as the random variable, and $F(1, 38) = 11.1$, with test words as the random variable. For the Globally Inconsistent texts, response times for the goal words were actually slower with the Problem text than the Control text, but this difference was not significant, $F(1, 49) = 1.31$ and $F(1, 38) = 2.58$.

The error rates for the goal test words were generally in accord with the response times. The interaction between Locally versus Globally Inconsistent text and continuation type was significant, with subjects as the random variable, $F(1, 49) = 4.21$, but not with test words as the random variable, $F(1, 38) = 2.37$. No other effects were significant ($Fs < 1.07$).

The reading time data is presented in Table 5 for completeness. There are two points worth noting. First, reading times for Sentence 6 are slow in all conditions, reflecting the point at which a new goal is introduced. However, reading times show less slowing for the Locally Inconsistent Problem continuations than for the other three conditions. This suggests that connecting a new goal to a previously mentioned higher order goal may be easier when the new goal is perceived to be directly related to previously mentioned goals. Second, the patterns of reading times are about the same for the two kinds of texts, Globally and Locally Inconsistent. Thus, the differences in response times for the goal test words cannot easily be ascribed to differences in reading times.

Discussion

Experiments 1 and 2 offer three tests of the notion that causal global inferences are encoded during reading. In both the Try Again and the Substitution conditions of Experiment 1, the general (global) goal should have been tied into comprehension at the ends of the stories. The same is true for the Globally Inconsistent Problem continuations of Experiment 2. However, in none of the three cases was there evidence that the general goal was more available after these continuations than after the Control continuations. Instead, the results support the hypothesis that global information is not automatically used during local comprehension.

Experiments 1 and 2 also offer two tests of the idea that the availability of concepts depends on whether they are required to establish local coherence. In both the Try Again continuations of Experiment 1 and the Locally Inconsistent Problem continuations of Experiment 2, concepts that were required for local coherence showed facilitation relative to the Control condition.

Experiments 1 and 2 used an on-line testing procedure. It is often argued that there are several possible interpretations of results obtained with this procedure (cf. McKoon & Ratcliff, 1980a; McKoon & Ratcliff, 1986; McKoon & Ratcliff, 1989a; Potts, Keenan, & Golding, 1988; Ratcliff & McKoon, 1988). Up to this point, we have assumed that a response to a test word reflects the state of availability of the concept tested, that is, the state of availability at the end of the text that precedes the test word. However, another possibility is that the response reflects a backwards context-checking process by which the test word is matched against the preceding text to determine if it fits the context (Forster, 1981). A poor match could inhibit the response, and a good match could facilitate it. Still another possibility is that the preceding text and the test word are jointly matched against memory as a compound cue (Ratcliff & McKoon, 1988); again, a good match would facilitate the response and a poor match would inhibit it. Fortunately, the data for Experiment 1 provide the means to decide among the interpretations. Both the backwards context checking and the joint matching interpretations lead to the same prediction: Response times for the general goal test words should be facilitated in the Try Again and Substitution conditions relative to the Control condition. This is because the texts in the former two conditions are still discussing information relevant to the general goal, whereas the Control is not (in the Control continuation, a new general goal has been introduced). For example, the test word *clean* should have provided a good context-checking match when the Try Again and Substitution continuations discussed water or brooms, but not when the Control continuation discussed painting a barn. However, this prediction does not fit the data; there were no significant differences in response times across conditions for the general goal test word. By this reasoning, we interpret the results of Experiments 1 and 2 as reflecting inference processes that occur during reading. The processes of backwards context checking and jointly matching text and test word against memory may also have been part of the processing of the test word, but they were not responsible for differential response times and accuracy rates across experimental conditions. However, it should be stressed that this is not a general conclusion about the on-line processing of test words. In other experiments, backwards context checking or a joint matching process might be responsible for on-line testing results.

The results of Experiments 1 and 2 support the local coherence, minimalist hypothesis over global inference theories. A recent experiment by Suh and Trabasso (1988) also can be interpreted to support the minimalist hypothesis (although Suh and Trabasso interpreted their results differently). They tested for the use of global information during reading of texts like that in Table 1 and found increased availability of global information at places that might have corresponded to coherence breaks, that is, points at which local coherence may not have been possible without the use of global information.

Despite the support for the minimalist hypothesis in Experiments 1 and 2, it could be argued that the texts in all these experiments were short and unnaturalistic. Also, only one experimental methodology was used, testing single word recognition immediately after reading. In Experiments 3 and 4, longer and more natural texts were used. The procedure in Experiments 3 and 4 was one that would allow examination of possible global inferences in the memory representations of the stories.

Experiment 3

The stories for Experiment 3 were 600-word narratives of the sort that might describe a television adventure story (see Tables 6 and 7). They were written to express a series of goals for a main character, with each goal eventually being fulfilled through some outcome. The goals were embedded such that fulfillment of any goal required that all of its subordinate goals had to be fulfilled first. For example, in the *Kidnapped* story, Jon had to help Ali with the microfilm to get into the fortress, and he had to get into the fortress to find his daughter, and so on. Once the most subordinate goal was fulfilled (e.g., Jon gets the microfilm), then the other higher goals could each be fulfilled in turn. If global causal inferences are constructed during reading, then each goal should be connected to its eventual outcome by inferred relations. This should be true even though the goal and the outcome events are far from each other in the text. However, if only local relations are constructed, then the goals will not be connected directly to their outcomes.

Whether the goals of the stories were connected to their outcomes in the encoded representations of the stories was tested with a priming procedure. Subjects read two stories and then were presented with a list of test statements for verification. For each story, there were statements that tested goals and statements that tested their outcomes. Theories that assume the encoding of global causal relations during reading would predict that a goal was connected to its outcome during reading and therefore that the connection would be encoded into the memory representation of the story. It follows, then, that a test statement about the goal should facilitate responses to an immediately following test statement about the outcome. This should be true even when several paragraphs intervene between the statements in the text. The facilitation given to the outcome statement by the goal statement should be greater than any facilitation that might be given by some other statement that was equally far away in the text.

Method

Materials. Twelve stories were written, each with a series of embedded goals. An example story is shown in Table 6, and the structure of the goals used in the experiment is shown in Table 7. (Table 7 does not represent the complete goal structure for all the goals for all the characters, only those goals relevant to the test conditions used in the experiment.) The stories were written so that each subgoal had to be fulfilled before the next highest subgoal could be attempted. So, for example, Jon had to find his daughter before he could attempt to rescue her. For each story, there was a series of true–false test sentences. One of these, an outcome target, expressed the outcome of one of the goals; for example, "Ali drove with Jon hidden in the trunk" expressed the means by which Jon achieved the goal of entering the fortress. A second, goal prime, test sentence expressed the goal ("Jon had to find help to get into the fortress"). A third test sentence (action near goal prime), a control condition, expressed some action that was near to the goal in terms of number of words in the story but not directly related to the goal; and a fourth sentence (near prime), another control condition, expressed an action that was near to the outcome in terms of number of words. Four more test sentences represented the same four conditions (goal, outcome, and two controls) with a different goal of the story. Finally, there were eight other sentences used as fillers in the test lists, three true sentences and five false sentences. The stories ranged from 579 to 613 words in length and from 53 to 59 lines when presented on a CRT screen. Each story was divided into seven paragraphs. The test sentences that represented the experimental conditions ranged from 7 to 11 words in length. Test sentences were taken as exactly verbatim from the stories as possible, allowing for shortening and using names or descriptions instead of anaphors.

There were also 12 other stories that were part of another experiment (Ratcliff & McKoon, 1988). These were about the same length, and each of them had seven true and five false test sentences that were used in the test lists.

Procedure. The experiment was conducted with a CRT screen and keyboard as in Experiments 1 and 2. The experiment began with a practice list of 40 strings of letters presented for lexical decision to give subjects practice at responding quickly and accurately with the keys on the CRT keyboard. After the lexical decision, there was 1 study–test list for practice and then 12 study–test lists for the experiment proper.

Each study-test list began with an instruction to press the space bar

Table 6
An Example Story From Experiment 3: Kidnapped

Jon was a CIA agent who often worked behind the Iron Curtain. He had made many enemies, and one of them, a KGB agent, kidnapped his daughter, Karyn, while she was on a trip to the Bahamas. It was all part of a plan to get revenge because Jon had foiled one of the enemy agent's plots many years before.

Jon wanted to get Karyn back from the KGB agent who had kidnapped her as quickly as possible. He had worked against the KGB agent, Vladimir, many years ago and was very worried about his daughter's safety. Although the authorities told Jon that he should stay at home and let the professionals do their job, Jon decided that he had to get to the Bahamas. Anxious, as any father would be, he made a reservation on the first plane he could find. In a few hours Jon arrived in the Bahamas.

Jon believed the only way he would get Karyn back safely was to find her himself. He had to find out where the kidnapper had taken her. Soon after Jon checked into a hotel, a young man delivered a ransom note from his enemy, Vladimir. As soon as the messenger left, Jon quietly followed him. He hoped that the young man would lead him to Vladimir. After some time, Jon arrived at a large, old fortress that was once used as a prison. As Jon watched the messenger go into the fortress, he was sure this was where his daughter was being held by Vladimir. He hoped that she was alright.

The fortress appeared to be completely impenetrable. Jon knew that if he was to rescue Karyn he would have to find help getting into it. Jon returned back to town, hoping to find a mercenary to help him. After visiting several bars, Jon met an old friend, Ali Al-Dib, a double agent he had known for many years. They had worked both against and with each other, but they always remained friends. Jon and Ali had some beers and talked over old times. Jon discovered that Ali had done business with Vladimir on several occasions. Jon explained his situation to Ali and asked him to help rescue Karyn.

Ali was busy with his own mission, stealing some microfilm that contained the locations of missile silos of certain west European countries. He hoped to sell it to the highest bidder. Ali agreed to help Jon if Jon would help him first. Jon thought it was a fair exchange and agreed to the bargain. They sat up late that night trying to come up with a plan to get the microfilm, which was hidden in the British embassy. They came up with a deceptively simple plan. Since Jon knew some people at the embassy, he would go in first and keep them occupied while Ali stole the microfilm. It worked.

Ali contacted Vladimir and asked him if he would be interested in buying the microfilm. Vladimir wanted to see it first, so Ali drove to the fortress with Jon hidden in the trunk. The guards recognized Ali and let him into the fortress without searching his car, so they did not find Jon in the trunk. While Ali kept Vladimir busy examining the microfilm, Jon ran from room to room and finally found the room where his daughter was being held hostage.

They escaped, undetected, and hid in Ali's car. Soon, Ali finished his business with Vladimir and got into the car. He drove Jon and Karyn to the airport before Vladimir realized Karyn had been rescued. In just a few hours, Jon and Karyn were safely back home.

Test sentences
- Outcome target: *Ali drove with Jon hidden in the trunk.*
- Goal prime: *Jon had to find help to get into the fortress.*
- Action near goal prime: *Jon met an old friend who was a double agent.*
- Near prime: *Jon kept the people at the embassy occupied.*

on the CRT keyboard. When the space bar was pressed, there was a 500-ms pause, and then the first paragraph of the first story was displayed. The paragraph remained on the screen until the subject pressed the space bar again; then the screen was erased, and after a 100-ms pause, the next paragraph was presented. Presentation continued in this way through all the paragraphs of the story. After the last paragraph, there was a 3-s pause, and then the second story was presented in the same way. After a 3-s pause after the second story, a row of asterisks was displayed for 500 ms, and then the test sentences were presented one at a time. Each sentence remained on the screen until the subject pressed a response key ("?/" for true and *z* for false), and then the screen was erased and there was a 100-ms pause. If the response was correct, the next sentence was presented immediately. If the response was incorrect, the letters of the word *ERROR!!* were displayed one at a time for 600 ms each. Then the screen was erased and the next sentence presented. After all 24 sentences of the test list, the instruction to press the space bar for the next study list was presented.

Table 7
Goal and Outcome Structure Kidnapped

Goal and structure
Goal 1: Rescue his daughter
Goal 2: Find his daughter himself
Goal 3: Get into the fortress
Goal 4: Help Ali with the microfilm
Outcome 4: Got the microfilm
Outcome 3: Got into the fortress
Outcome 2: Found his daughter
Outcome 1: Escaped with his daughter

The stories presented in each of the 12 lists were chosen randomly, except that there was one story from the experiment (and one from the other experiment) in each list. These two stories were presented in random order. The test sentences of a list were presented in random order (sentences from the two stories interspersed), except for two restrictions: The test sentences used in the experimental design were not presented in the first test position, and the test sentence immediately preceding a prime–target pair was not from the same story as the target. A different randomization was used for every second subject.

Subjects and design. There were two groups of subjects. For the first group (21 subjects), the outcome test sentence was primed by its goal test sentence, the test sentence near to it in the text, or a test sentence from the other story of the study list (Control). These three conditions were combined in a Latin square design with sets of subjects (7 per set) and sets of outcome test sentences (8 per set). For the second group of subjects (32 subjects), an outcome test sentence was primed by its goal, another action near to the goal, or a test sentence from the other story (Control). Again, the three conditions were combined in a Latin square with sets of subjects and sets of outcome test sentences. The subjects participated in the experiment for credit in an introductory psychology course.

Results

Means were calculated for each subject and test sentence in each condition, and means of these means are shown in Table 8. For the target test sentences, only responses preceded by a correct response to the priming sentence are included in the means.

The minimalist prediction is that responses to the outcome targets should receive the largest amount of facilitation when the prime is the sentence near to the outcome in the text. There should also be some facilitation when the prime is a sentence farther away in the text (because the sentences are from the same text), but the amount of this facilitation should not depend on whether the prime is related to the target as goal and outcome. This is the pattern of data shown in Table 8. Relative to the Control condition (the prime from another story), responses to the outcome target are fastest with the near prime and about equally fast with the Goal and Action Near the Goal primes.

For the first group of subjects, an ANOVA showed that the overall difference in response times for the target test sentences was significant, $F(2, 40) = 19.0$, with subjects as the random variable, and $F(2, 22) = 17.6$, with test sentences as the random variable. Post hoc tests showed that response times in the near priming condition were faster than response times in the goal priming condition, $F(1, 40) = 4.6$, and $F(1, 22) = 5.3$. Standard error of the response time means was 38 ms. There were no significant differences in error rates.

For the second group of subjects, an ANOVA also showed that the overall difference in response times for the target test sentences was significant, $F(2, 62) = 18.3$, with subjects as the random variable, and $F(2, 22) = 15.2$, with test sentences as the random variable. Post hoc tests showed that the difference between the goal and action near goal priming conditions was not significant ($Fs < 1.0$). The standard error of the means was 30 ms. Differences in error rates were not significant ($Fs < 1.3$).

For the first group of subjects, the mean reading time per paragraph was 12.150 s, and for the second group of subjects, it was 12.939 s.

Table 8
Results From Experiment 3: Response Times (in Milliseconds) and Error Rates for Outcome Target Sentences and Filler Test Sentences

	Subject Group 1		Subject Group 2	
Priming condition	RT	% error	RT	% error
	Outcome target sentences			
Goal prime	1,567	6	1,541	6
Action near goal prime			1,576	7
Near prime	1,451	7		
Control prime	1,781	11	1,772	10
	Filler test sentences			
True items	1,605	8	1,628	10
False items	1,801	15	1,823	26

Note. RT = response time.

Discussion

If global inferences connected goals to outcomes in the stories of Experiment 3, then the outcome test statements should have been primed more by the goal test statements than by the action near goal test statements. However, the two priming effects were not significantly different. Once again, as with Experiments 1 and 2, the data failed to provide evidence of global inferences.

One problem that might be raised with Experiment 3 is that the data show no evidence of any kind of structure for the stories at all. Responses to the target statements were facilitated more by other statements from the same story than by statements from a different story, but within a story the only effect was one of surface distance, with the near primes giving more facilitation than the other within-story primes. However, as discussed earlier, previous investigations of mental representations of texts have demonstrated some internal structure, specifically, that propositions sharing arguments are connected together (McKoon, 1977; McKoon & Ratcliff, 1980b; Ratcliff & McKoon, 1978). In Experiment 4, we looked for evidence of this kind of structure.

Experiment 4

If propositions from the stories of Experiment 3 are connected by argument repetition during reading, then evidence of those connections should be observable in priming effects. For example, all the propositions about Ali should be connected together, whether he was explicitly called Ali or referred to as "double agent" (cf. McKoon & Ratcliff, 1980b). These propositions should be more closely connected to each other than they are to other propositions that do not refer directly to Ali. We tested for these differences in connections with the same procedure as in Experiment 3, priming in verification of statements from the stories.

Method

Materials. The 12 stories from Experiment 3 were used, with a new set of test sentences. For each story, there were two target test sentences. Each of these targets had two primes. One of the primes was near the target in terms of the argument repetition structure of the story, and the other was relatively far from the target. The average distance of the two primes from the target in terms of number of words was about the same (191 words and 192 words, respectively). For example, in the *Kidnapped* story, one target was "Jon met Ali, who was an old friend." The near prime for this target was "Karyn's father took the first plane to the Bahamas," which shares an argument with the target because Jon and Karyn's father are the same person. The far prime for this target was "Vladimir wanted to see the microfilm before he bought it," not so closely connected to the target by argument repetition. The number of words in the prime and target test sentences ranged from 7 to 11. There were also 8 filler test sentences for each story, 3 true sentences and 5 false sentences.

Procedure. The procedure was the same as for Experiment 3, except there were no stories from another experiment so that the total number of study-test lists was six.

Design and subjects. Each target was primed by another test sentence near it in argument repetition structure, another test sentence far from it in argument repetition structure, or a sentence from the other story in the study list (Control). These three conditions were combined in a Latin square with the 12 stories (4 per set) and 24 subjects. The subjects participated for credit in an introductory psychology course.

Results

The data were analyzed as in Experiment 3, and the results are shown in Table 9.

As expected, response times for the targets were speeded with the near prime, relative to both the prime from the other story and the far prime from the same story. An ANOVA showed that overall differences were significant, $F(2, 46) = 8.5$, with subjects as the random variable, and $F(2, 22) = 5.9$, with test sentences as the random variable. The difference between the near and far conditions was significant by post hoc test, $F(1, 46) = 4.6$ and $F(1, 22) = 6.7$. The standard error of the means was 26 ms. There were no significant differences in error rates (Fs < 1.9). The mean reading time for all paragraphs was 17.260 s.

Discussion

The motivation for Experiment 4 lay in a potential problem with interpretation of the results of Experiment 3. We want to claim that, for the stories of Experiment 3, readers encoded the same local relations as have been demonstrated in past experiments. The inferences that they failed to encode were the global ones for which we tested. However, Experiment 3 gave no evidence that readers had, in fact, encoded any relations at all other than proximity in surface distance. Experiment 4 provided this evidence, showing that relations based on argument repetition were represented in memory. Thus, the mental representation does show structure, but the structure is based on argument repetition and not on global inferences about causality.

Table 9
Results From Experiment 4: Response times (in Milliseconds) and Error Rates for Target Test Sentences and Filler Test Sentences

Priming condition	RT	% error
Target test sentences		
Near prime	1,502	5
Far prime	1,579	5
Control prime	1,651	10
Filler test sentences		
True items	1,586	10
False items	1,661	21

According to the minimalist hypothesis, the inferences that build the argument repetition structure are based on information that is easily available, in this case, the names and descriptions of the characters in the stories. For example, Jon, the CIA agent, is the main character in the story in Table 6. Whenever Jon is mentioned in the story, and new propositions are to be attributed to him, his name serves to make available other information encoded about him earlier in the story and to make it likely that these different pieces of information will be connected through repetition of their argument *Jon.* So long as a definite description of an entity is a strong enough cue to evoke previous information about the entity, then the different pieces of information can be connected together.

An argument that might be advanced against the minimalist interpretation of the results of Experiments 3 and 4 is that a recognition test procedure does not tap the level of representation at which inferences are encoded, but instead some more superficial level of representation. However, this argument is countered by previous research. First, discussed later, recognition does give evidence for some kinds of elaborative inferences (those supported by well-known, easily available information). Second, recognition also gives evidence for structural inferences when the minimalist hypothesis predicts that such inferences should be encoded. McKoon and Ratcliff (1980b) used recognition to show that the organization of a list of sentences was inferred from well-known (schema) knowledge. Similarly, McKoon, Ratcliff, and Seifert (1989) used recognition to show that the relations between stories were inferred from schema knowledge. Recognition can also be used to show that both structural and elaborative inferences are constructed when subjects are given instructions to use special strategies during reading (Seifert, McKoon, Abelson, & Ratcliff, 1986; M. McDaniel, November, 1991, personal communication).

In sum, Experiments 1 through 4 strongly support the minimalist hypothesis over the constructionist hypothesis. With both simplistic and natural texts and with both on-line and delayed memory procedures, there was no evidence that causal global inferences were constructed. Evidence for global inferences appeared only for texts that were not locally coherent. These results emphasize a striking contrast between local and global inferences. Local inferences have been easy to demonstrate empirically in a large number of studies. However, in the same kinds of experiments in the same laboratory situations, there is no evidence for the kinds of causal global inferences posited by a number of theorists.

It is important to recognize that the results of Experiments 1 through 4 demonstrate failures to encode global—not local—causal inferences. The minimalist claim is that local causal inferences will be encoded either if they are easily available from long-term memory or if they are required to establish local coherence.

Van den Broek (1990) and Fletcher and Bloom (1988) have proposed a model by which the causal inferences necessary for local coherence are encoded. The architecture and processes of the model are the same as in van Dijk and Kintsch's (1983) model, except that the propositions of a text are connected by causal relations in addition to argument-repetition relations. The model has the same short-term memory limit on processing as the minimalist position: Only propositions that are in short-term memory at the same time are connected by inferences; information from other parts of the text is used only if the local information is not coherent. Van den Broek provided a definition for *coherence* in terms of four criteria of causality. Coherence is maintained for an event if there are antecedents for the event that are temporally prior, operating at the time of the event, necessary for the event to occur, and sufficient for the event to occur. The event is connected to antecedents that fulfill these criteria just as propositions containing the same argument are connected in the Kintsch and van Dijk model. Only if there is no antecedent fulfilling all the criteria does a coherence break occur (van den Broek, 1990, p. 434). Then, either propositions of the text that are no longer in short-term memory are retrieved, or new propositions are generated to provide the connections necessary for coherence. Evidence consistent with this model has been provided by Bloom et al. (1990) and by Fletcher and Bloom (1988).

The results of Experiments 1 through 4 show that the global causal inferences defined by recent theories are not part of automatic encoding processes. However, the results say nothing about their roles in other more goal-driven encoding processes or in retrieval processes. Although the focus of this article is on inferences that are constructed automatically during reading, it must be stressed that understanding the processes that construct inferences important to a reader's goals and the processes underlying recall are also extremely important. Practically speaking, we use goal-driven reading processes and recall processes ubiquitously, and setting up optimal reading and recall processes is the aim of many educational efforts. The problem raised by the results presented in this article is to accommodate a minimalist representation of textual information with the more constructionist information that appears in recall and question answering and that readers use in those frequently occurring situations where they have specific goals. One possibility is that information beyond the minimal is constructed by retrieval processes that follow local connections through memory. A model like this, based on Raaijmakers and Shiffrin's (1981) recall process, has been developed by Fletcher and van den Broek (1989), with some empirical support. In general, however, there is little current theorizing about the more strategic aspects of text processing.

Elaborative Inferences

The most important claim of many mental models theories of text comprehension is that the mental representation of a text automatically depicts the events described by the text in a lifelike way. Various parts of the description must be constructed by elaborative inferences, because a text seldom provides an explicit description of an event that is sufficiently complete to describe the situation in a lifelike way. Thus, it is essential to mental models theories to show that elaborative inferences are automatically encoded during reading.

In contrast, the minimalist hypothesis does not make any claim about the extent to which a mental representation depicts the event described by a text. Instead, the minimalist hypothesis applies other criteria to decide whether inferences will be constructed: whether the text is locally coherent and whether the information necessary for an inference is easily available. Usually, these criteria are not consistent with a full description of a textual event. This is because the information necessary for a complete description is usually not all easily available. Also, for local information, a coherent description is not necessarily a complete description.

The minimalist criteria for elaborative inference processes are advantageous in that they provide guides to empirical research. Specifically, demonstrations that a criterion for elaborative inference is met (e.g., a demonstration that inference-supporting information is quickly available; McKoon & Ratcliff, 1989b, Experiment 2) can be separated from demonstrations that elaborative inferences are encoded, thus avoiding circularity. The criterion provided for elaborative inferences by a constructionist hypothesis does not so obviously lead to independence: There is no a priori way to know, for any particular inference, whether it is required in the representation of an event. As a result, there is no way to independently verify whether a particular inference should be encoded.

In the sections that follow, specific kinds of elaborative inferences are considered. Each case allows evaluation of the minimalist hypothesis, the constructionist hypothesis, or both. According to the minimalist hypothesis, for each kind of inference, encoding should depend on the availability of the information necessary to support inference processes. If supporting information is not quickly available, then an inference should not be constructed (unless necessary for local coherence). According to the constructionist hypothesis, the encoding of inferences should not depend completely on the availability of supporting information; instead encoding should depend on whether an inference is required for a lifelike description of the event described by the text.

Consideration is limited to those kinds of elaborative inferences for which there is sufficient research to provide a reasonably coherent body of data. These are instrumental inferences, inferences about the meanings of words, and predictive inferences about what will happen next in a story. For other elaborative inferences, such as expectations (Duffy, 1986), and inferences deriving from the argument structures of verbs (Boland, Tanenhaus, & Garnsey, 1990; McKoon & Ratcliff, 1989c; Tanenhaus, Carlson, & Trueswell, 1989), the accumulated data are not sufficiently constraining to test the minimalist and constructionist hypotheses.

Instrumental Inferences

When elaborative inferences were first studied extensively, in the 1970s, it was argued that a description of the event described by "Mary stirred her coffee" (Dosher & Corbett, 1982) should include the instrument *spoon* (cf. Johnson et al., 1973; Paris & Lindauer, 1976). Early evidence to support the encoding of instrumental inferences came from cued-recall studies, in which recall of a text was facilitated by a cue that was an instrument highly associated with a verb in the text but not stated explicitly in the text (Paris & Lindauer, 1976). Subsequently, Singer (1978, 1979) and Corbett and Dosher (1978) showed that cued-recall results could not decide issues of encoding.

More recent research argues against the constructionist hypothesis. Dosher and Corbett (1982) looked at the relation between an inference sentence and its implicit instrument, for example, "Mary stirred her coffee" and *spoon*. They examined whether the relation would affect responses to the instrument when it was presented as a test item in a Stroop task. Results showed that Stroop responses were not affected. There was no effect regardless of whether the instruments were the most likely for their sentences, and there was no effect for instruments that were tools or for instruments that were body parts. Only when subjects were instructed to explicitly guess the instrument in advance of the Stroop test were responses affected. In other words, unless an instrument was explicitly requested, there was no evidence that it was involved in comprehension of the inference sentence. These results argue strongly against the constructionist hypothesis because a complete description of an event like stirring coffee seems to require an instrument.

On the other hand, the results are compatible with the minimalist hypothesis, given the assumption that the instruments were not automatically available during reading of the inference sentence. The assumption can be tested as follows: If the availability of the instruments is increased to a sufficiently high level, then they should be encoded. This test was part of a study by McKoon and Ratcliff (1981). Availability was increased by explicitly mentioning an instrument several sentences before the inference sentence for which it would be the implicit (but highly typical) instrument; for example, spoon would be mentioned several sentences before the sentence "Mary stirred her coffee." The instrument was presented as a test word immediately after the inference sentence. Responses to the test word were facilitated (relative to a control condition), suggesting that the relation between sentence and instrument was available in an immediate test situation. This availability should lead to encoding according to the minimalist hypothesis. That the instrument was encoded was confirmed by a priming effect in a delayed memory test. Presenting the instrument as a test word immediately before a noun from the inference sentence (e.g., *spoon* immediately before *coffee*) facilitated responses to the noun (relative to a control condition). The facilitation indicates a close association between the instrument and the noun in the memory representation of the sentence, which in turn indicates that the instrument was encoded with the sentence.

Overall, empirical results from studies of instrumental inferences favor the minimalist hypothesis. Highly typical instruments of verbs are strong candidates for inclusion in a mental model of a stereotypical event such as stirring coffee, yet there is no evidence that they are used in comprehension or that they are encoded (unless subjects engage in special strategies; Dosher & Corbett, 1982). In contrast, the minimalist hypothesis predicts the finding that increasing the availability of the instruments during comprehension leads to their encoding.

Inferences About the Meanings of Words

Instrumental inferences were one of the kinds of elaborative inferences studied in the 1970s in the effort to document constructed mental representations. Another kind were inferences about the meanings of words. For example, R. C. Anderson and Ortony (1975) used cued recall to examine the meaning of *container* in the sentence "The container held the apples." The cue for the sentence was either *basket* or *bottle*, and *basket* was more effective.

Results like this suggest that contextually appropriate aspects of the meanings of words might be encoded into the mental representations of texts, and current research confirms this idea. McKoon and Ratcliff (1988; see also Barsalou, 1982; Tabossi, 1982; Tabossi & Johnson-Laird, 1980) used texts in which a specific feature of a noun was made salient (e.g., a text about painting a picture of a tomato should make salient the color of the tomato, red). After a series of texts, test sentences were presented for verification. Sentences that tested a feature that had been made salient in the text (e.g., "tomatoes are red") were verified faster than control sentences.

If features of the meanings of words are automatically encoded into memory, then according to the minimalist hypothesis they must have been easily available during comprehension. Easy availability should show up when the features are tested immediately after reading. Immediate facilitation of this sort was obtained by McKoon and Ratcliff (1988; see also Tabossi, 1982; Tabossi & Johnson-Laird, 1980), using a sentence verification task, and by Greenspan (1986) using lexical decision.

Research that has examined the contextually defined meanings of category terms is also consistent with the minimalist hypothesis. If a text mentions the category *animals* in the context of milking some animals on a farm, subjects have difficulty in later rejecting the word *cow* as having appeared in the text (McKoon & Ratcliff, 1989b). This result can be taken to indicate that something like the concept *cow* was encoded into the mental representation of the text. It should follow that the concept *cow* is easily available during comprehension. This availability appears as facilitation when *cow* is tested immediately after the text, with both recognition and lexical decision (McKoon & Ratcliff, 1989b), and also as faster reading time for a follow-up sentence that explicitly mentions *cow* (Roth & Shoben, 1983).

These patterns of results are consistent with both the minimalist and constructionist views, but the minimalist view is the more constrained. The hypothesis that encoded inferences must be based on immediately available information would be contradicted if some inference was encoded, but its supporting information was not quickly available during reading (and subjects did not engage in special strategies). However, there are no inferences that pattern this way. In contrast, the constructionist view makes no claims about the relation between availability during reading and subsequent encoding, and so no constraints are placed on the constructionist hypothesis.

Predictable Events

If someone falls off a 14-story roof, then the real-life result will be death. Because the outcome is so predictable, a mental model for a text such as "the actress fell from the fourteenth story" should automatically include the inference that she died. It would not be reasonable, from the mental model point of view, to leave her suspended in midair. On the other hand, the inference about death is not necessary for local coherence if the text ends with the sentence about the fall. The event of falling from a 14-story building is not familiar enough to make the inference easily available. So the minimalist hypothesis predicts that the inference about death will not be included automatically in the mental representation.

To test for inferences of this kind, McKoon and Ratcliff (1986, 1989d, 1989e) used a speeded recognition memory test. Subjects read several short texts before reading a list of test words. Each test word was followed by a signal, and the subjects were instructed to give a response immediately when the signal was presented. The delay between test word and signal was short enough that slow, strategic processes (that might construct inferences at the time of the test) were eliminated. The critical test words were those that represented inferences about predictable events where the events were known to be highly predictable from previous norming studies. For the actress text, the critical test word was *dead.* The correct response for these test words was *no,* because they had not been explicitly stated in any of the studied texts. However, if the inference was generated during reading, then a negative response should be difficult and subjects should tend to make errors (relative to a control condition, in which the subjects read a text that did not predict the critical event).

When a critical word was presented for test, it was preceded by a priming word (displayed for 200 ms). In one condition, the priming word was the neutral word *ready.* In this condition, subjects did not make significantly more errors when they had read the text that predicted the critical word than when they had read the control text. This result indicates that the predictable event was not clearly and explicitly encoded during reading of the predicting text, counter to the constructionist hypothesis. However, in a second condition, the prime for the critical test word was a word from the text (e.g., the word *actress*). In this condition, subjects did make more errors when they had studied the predicting text relative to the control text.

McKoon and Ratcliff (1986, 1989a, 1989d, 1989e, 1990; Potts et al., 1988) interpreted this increase in errors with the prime from the text as evidence for partial encoding of the inferences. Although the failure to find an elevated error rate with the neutral prime indicates that the inference could not have been explicitly and completely encoded, the increase in error rate with the prime from the text indicates that the inference was encoded to some degree. On the basis of this result, McKoon and Ratcliff suggested that inferences were encoded to varying degrees, with some inferences encoded minimally by a set of features or propositions that do not completely instantiate the inference. This proposal is supported by findings that inferences are encoded to a higher degree if they are based on well-known information, such as semantic associations or category membership (McKoon & Ratcliff, 1989b, 1989d).

If inferences about predictable events are not explicitly encoded, then according to the minimalist position, the reason should be that they are not quickly and easily available during reading. Several experiments have shown that this is the case (see McKoon & Ratcliff, 1989e). When the textual information that would generate the inference is immediately followed by a test for the inference, then responses on the test are not affected (McKoon & Ratcliff, 1989d; Till, Mross, & Kintsch, 1988). For example, in the sentence, "The diver jumped, spun, and hit the cement," the information necessary to know that he was hurt is given only by the final word of the sentence. When the test word *hurt* was presented immediately after the final word of the sentence, subjects had no difficulty in deciding that it had not appeared in the text (relative to a control condition). However, when more time (and/or other textual material) intervenes between text and test, then the inference does affect responses (McKoon & Ratcliff, 1986; Potts et al., 1988; Till et al., 1988). In contrast, when well-known information from general knowledge is available to support an inference about a predictable event (e.g., the predictable event of sitting after approaching a chair), then the inference does affect responses to a test word, and it does so even when the test word is presented immediately after the textual information that would generate the inference (McKoon & Ratcliff, 1989d). This contrast, between those predictable event inferences that are supported by well-known information and those that are not, is exactly in accord with the minimalist hypothesis.

The results on predictable inferences, like the results on instrumental inferences, disconfirm the constructionist hypothesis: Inferences that should be explicitly represented in a mental model, like the death of the actress, are not. However, the results on predictable inferences are also not consistent with an all-or-none minimalist position. Instead, the results suggest that inferences vary in the degree to which they are encoded. This suggestion is taken up in the General Discussion section.

Inferences From Situation Models

A number of different theories embrace the constructionist hypothesis. Theories proposed to explain story understanding hypothesize that readers construct connections between different parts of a story, such as goals and outcomes (Mandler, 1978; Mandler & Johnson, 1977; Rumelhart, 1975, 1977; Stein & Glenn, 1979; Trabasso & van den Broek, 1985). Theories proposed to explain the understanding of descriptions of events assume that the mental representation is "filled out" with inferred information (Bower, Black, & Turner, 1979; Glenberg et al., 1987; Johnson-Laird, 1980; Morrow et al., 1989; van Dijk & Kintsch, 1983). The hypothesis that unifies these theories is that the mental representation of a text automatically specifies, in some complete way, the real-life situation described by the text. The mental representations are labeled *mental models* or *situation models.*

These terms, *situation model* or *mental model,* do not in principle have to incorporate elaborative inferences beyond those postulated by the minimalist position. For example, a situation model might be proposed that contains only those elaborative inferences that are easily available from general knowledge. These inferences might connect propositions of the text in ways that simple argument repetition would not, relating the propositions to reflect well-known knowledge. Such a model has been proposed by Kintsch (1988), and this model is discussed further in the General Discussion section. In this section, evidence pertinent to constructionist situation models is reviewed.

The constructionist hypothesis is that readers automatically construct a full representation of the real-life situation described by a text. This hypothesis has been tested directly in a number of experiments. These experiments differ from those that investigate elaborative inferences in that they use a situation rather than a text as their starting point. For the experiments on elaborative inferences, the issue of concern is the relation between a given text and the encoding of some specific inference. The issue of concern for experiments on situation-based inferences is the relation between the mental representation of a text and a real-life situation (or a lifelike situation learned in an experiment).

Many of the early experimental results thought to demonstrate the use of lifelike situation models during reading have since been reinterpreted. An excellent review of this work has been provided by Alba and Hasher (1983), and only several main points are repeated here. One kind of experiment used passages that were extremely difficult to understand and recall unless prior knowledge of the situation was invoked (e.g., the "washing clothes" passages used by Bransford & Johnson, 1972; see also Dooling & Lachman, 1971). It was originally claimed that making available prior knowledge of the situation led to the use of a full mental model during reading; however, prior knowledge may have simply provided a specific context for the interpretations of individual words and the construction of locally coherent structures, in accord with a minimalist approach (see also Alba et al., 1981). A second kind of experiment used short sentences that could be combined to describe an event (e.g., ants eating jelly on a kitchen table; Bransford & Franks, 1971). In a recognition test, subjects' confidence that they had studied a sentence describing the whole event was greater than their confidence that they had studied shorter sentences describing parts of the event, even though they had never studied the sentence describing the whole. However, it was later shown that this result could be obtained with meaningless material, such as nonsense syllables, suggesting that subjects had actively engaged in special encoding strategies (cf. J. R. Anderson & Bower, 1973; Flagg, 1976; Flagg & Reynolds, 1977; Katz & Gruenewald, 1974; Moeser, 1976; Reitman & Bower, 1973). Third, several experiments were thought to demonstrate the use of knowledge about prototypical situations (schemas) during reading (Bower et al., 1979; Graesser, 1981); for example, subjects were more likely to recognize a highly typical schema action as previously studied than a less typical action. However, Alba and Hasher pointed out that the effect can be explained as a response bias. Also, consistent with a minimalist position, it was later shown that the use of schema knowledge depends on how available it is during comprehension. Only when schema relations are extremely well-known are they automatically used to relate events during reading (McKoon et al., 1989; Seifer et al., 1986).

More recently, the lifelike situation models that have been tested empirically derive from theories proposed by van Dijk and Kintsch (1983) and by Morrow et al. (1987). The characteristics of these theories have been listed by Glenberg et al. (1987):

> A situation model is the result of interactions between information given in a text and knowledge about linguistics, pragmatics, and the real world; a situation model can be modified as new information comes in to produce a completely new interpretation of the text; the information in a situation model can be manipulated to produce emergent relations; a situation model is perceptual-like; a situation model guides interpretation of referential terms; and a situation model guides the generation of inferences (p. 69).

Since the 1970s and the realization that cued-recall experiments could not distinguish inferences generated at recall from inferences generated at encoding, there have been surprisingly few studies designed to investigate interactions between textual information and knowledge of lifelike situations. Some experiments (e.g., Johnson-Laird, 1980; Mani & Johnson-Laird, 1982; Perrig & Kintsch, 1985) used descriptions of situations (such as a textual description of the layout of a town; Perrig & Kintsch, 1985) but used procedures that invite subjects to engage in strategic processing (by extended study or a problem-solving type of task). Others have confounded learning instructions with situational versus other kinds of information (Schmalhofer & Glavanov, 1986). Only a few experiments have used procedures where exposure to a text is limited to one reading at an approximately normal reading rate.

Several of these experiments have been conducted by Morrow and colleagues (Morrow et al., 1987, 1989). They investigated whether knowledge of lifelike situations affects comprehension of narratives. With a map, subjects were taught about the rooms in a laboratory and the objects in those rooms. After the subjects memorized this information, they were presented with a series of narratives, each describing a character moving through the rooms. The subjects were interrupted at various points during the narratives with questions about whether two objects were located in the same room. Results showed that subjects were faster to answer the questions when the objects were located in a room that was relevant to the character's current location or the character's goal location.

There are two problems with taking these results as strong evidence against the minimalist hypothesis. First, subjects knew that they would be tested on the objects in the rooms as they read the narratives (all test questions were about pairs of objects). Subjects could plausibly adopt a strategy to perform well on the test questions (i.e., up to the level that would be expected of Stanford undergraduates); the strategy would be to rehearse objects while reading the narratives. At any point during reading, the probability of rehearsing the objects from a particular room could well depend on the room's relevance to the information being read at that point. If so, then the objects would be made available, not as the result of automatic (priming) processes but as the result of strategic retrieval processes by which a relevant room would be used as a retrieval cue for rehearsal of its objects. By this account, Greenspan et al.'s (1987) results are not due to the reader moving (metaphorically) through a situation model, complete with objects in their correct rooms, but instead to the reader's appreciation of the relative saliencies of concepts in local parts of the discourse and the use of the most salient concepts as retrieval cues. This account is consistent with the minimalist hypothesis, and support for it has recently been provided by Wilson, Rinck, McNamara, Bower, and Morrow (1992).

The second problem with these studies is that situation models do not predict which parts of a situation will be relevant to different narratives. Morrow et al. (1989) stated that, for the sentence "We flew from Paris to New York last week" (p. 300), comprehension is unlikely to involve information about the Atlantic ocean. This may be true, but then it becomes unclear why comprehension of a sentence about a character going from a conference room to a laboratory should make available (unstated) information about the shelves in a library that the character passes through on his way. In fact, one might argue that flying over the Atlantic ocean makes that ocean (and its perils) more salient than the shelves on the wall of a room that is quickly left behind. The problem is that there is no way of predicting what aspects of a situation are salient in any given situation and therefore no way of predicting which inferences should be included in the mental model.

This problem is critical for a constructionist situation model approach to discourse comprehension. If empirical data is to support the inclusion of lifelike information into a mental representation of discourse, then there must be clear, theoretically motivated distinctions between inferences that should be included in the situation model and inferences that should not be.

Experiment 5

One way to begin to define constructionist inferences is to consider the real-life situation that a text describes and assume that whatever information is in the real situation is also in the mental model. This is the approach taken by Glenberg et al. (1987). Subjects read short narratives like "A girl was enjoying the warm spring weather. She walked up to the entrance of a park, and bent down to pick up a flower for her sister. Then she walked into the park and down to a small stream where some ducks were feeding. She smiled to see seven tiny ducklings trailing behind their mother." If readers construct a situation model while reading this text, then at the end of the text, their model should include the girl, and the girl should have the flower with her, exactly as would be the case in real life. This model should be different from the model constructed for a second, control, version of the text. The control version was the same as the first version except that the girl bent down to smell the flower, she did not pick it or take it with her. In the model at the end of the control version, the girl would not have the flower with her. To test for the use of a situation model during reading, Glenberg et al. presented a recognition test word at the end of a text; for this example, the test word was *flower.* Glenberg et al. predicted correctly that responses to the test word would be facilitated when the girl had picked the flower to take with her compared with when she had only smelled it.

This result appears, at first, to provide elegant support for the notion that a situation model is used during comprehension. However, there is an alternative interpretation of the data. It might be that the differential response times to the test words result from their differential salience (or topicality) in a propositional representation. A flower picked to take with the girl for a present might be treated during comprehension as more relevant to the topic of its discourse than a flower smelled for a moment and then left behind (see discussions of discourse models by Grosz et al., 1983; Sidner, 1983a, 1983b; Webber, 1983). In Experiment 5, this alternative interpretation was tested by changing Glenberg et al.'s texts to add words that were topical but not model relevant. For example, the two versions of the flower text were changed so that the girl *bent down to an ornamental display* to pick a flower or smell a flower. The added propositions about the ornamental display contain the concept *flower,* and so they should vary in topicality as *flower* varies in topicality (cf. Kintsch, 1974). However, obviously, the display cannot accompany the girl, so the display cannot move with the girl in a situation model. Whether the girl picks the flower or merely smells it, the display is not part of the current situation at the end of the text. Thus, there are clearly contradictory predictions: In a situation model, when the girl picks the flower, the flower should be currently available at the end of the text but the ornamental display should not be. In a propositional representation, when flower is more salient, the display should also be more salient.

The procedure for Experiment 5 was the same as that used by Glenberg et al. (1987). Subjects read each text at a rate they controlled themselves, and, at some point during the text, a test word was presented for recognition.

Method

Materials. The 24 experimental texts were based on paragraphs used by Glenberg et al. (1987). For each text, there was a critical noun used by Glenberg et al. as the test word. In one version of the text, this noun stayed with the main character of the story as the action moved forward through the story. In the other version of the text, the noun was left behind the character as the action moved forward. We hypothesized that the critical noun was more salient in the texts for which it stayed with the main character than in the texts in which it was left behind. The texts were modified from those used by Glenberg et al. by the addition of a location for the critical noun. The location was always mentioned with the critical noun. For example, in the story just given as an example, the phrase "to an ornamental display" was added to give a location for the critical noun *flower.* The location was something that could not move with the main character.

The first sentence of each text was the same in both versions and served to introduce the main character, it averaged 10 words in length. The second sentence mentioned the critical noun and the location and was presented in one of two versions to manipulate whether or not the critical noun stayed with the character (a mean of 17 words in both cases). The third and fourth sentences completed the story (averaging 14 and 12 words, respectively). There was also a yes–no question associated with each text to test general comprehension of the story; the correct answer to 13 of the questions was yes, and the correct answer to 11 was no.

Filler texts (the same filler texts as were used by Glenberg et al., 1987) were chosen from a pool of 58 texts, ranging from 20 to 60 words in length. For 22 of these texts, there was a test word that had appeared in the text, and for the remainder, the test word had not appeared in any text. For 28 of the filler texts, the correct answer to the comprehension question was yes.

Procedure. The texts and test items were presented on a CRT screen, and responses were recorded on the CRT's keyboard. The CRT was controlled by a real-time microcomputer system.

The experiment began with a list of 30 lexical decision test items used to familiarize subjects with the response keys. After this practice, 10 filler texts were presented, and then the experiment proper began, with the 24 experimental texts and 24 filler texts randomly ordered. Presentation of each text began with the message "Press space bar" to initiate the text. When the space bar was pressed, the first sentence of the text was displayed. It remained on the CRT screen until the space bar was pressed again; then the screen was erased, and the next sentence was displayed. Sentences were presented in this way until the final sentence before the test word. When the space bar was pressed after reading of this sentence, the test word was displayed, with a row of asterisks underneath it. The test word remained on the screen until a response key was pressed, "?/" for a positive response if the test word had appeared in the text just read or *z* key for a negative response if the test word had not appeared in the text. If the response was not correct, then the word *ERROR* was displayed for 1,500 ms before the next sentence or yes–no question. If the text was one of the experimental texts, the test word was presented after the third sentence, and the fourth sentence was presented after the test word. For the filler texts, the test word was always presented after the last sentence of the text. After the text and its test word, the yes–no question appeared, and it remained on the screen until a response key was pressed. If the response was correct, the message to initiate the next text was presented. If the response was an error, then *ERROR* was displayed for 1,500 ms.

Subjects and design. There were four experimental conditions: The critical object either stayed with the main character until the end of the text remained behind, and the test word was either the critical noun or its location. These four conditions were combined in a Latin square design, with four groups of subjects (9 per group) and four groups of texts (6 per group). The subjects were 36 undergraduates from the same population as in Experiments 1 through 4.

Results

Means were calculated over responses for each subject and item in each condition, and means of these means are shown in Table 10.

Glenberg et al. 1987 showed that response times for the test word that was the critical noun were faster when the noun stayed with the character than when it was left behind. Table 10 shows a clear replication of this result; responses times for the critical noun test word were 66 ms faster when the noun stayed with the character. The important question is whether this same result obtains for the location test word. Our hypothesis was that the speed up in response times for the critical noun was due to its increased salience, not to the fact that it stays with the character. If this hypothesis is correct, then the speed up should also be obtained for the noun's location. The data confirm this hypothesis: Location response times were 55 ms faster when the critical noun was more salient (i.e., when the noun stayed with the character). This speed up in response times for the location test words is not predictable from a situation model. Thus, the results support the hypothesis that it was salience, not availability in a situation model, that was responsible for Glenberg et al.'s finding.

An ANOVA showed significantly faster response times for both noun and location test words when the noun stayed with the character, $F(1, 35) = 6.9$, and $F(1, 20) = 9.2$, and marginally significantly faster response times for the noun test words than the location test words, $F(1, 35) = 5.5$, and $F(1, 20) = 2.9$. The interaction between the two factors was not significant ($Fs < 1$). The standard error of the response time means was 18 ms. There were significantly more errors on the location test words than the noun test words, $F(1, 35) = 20.4$, and $F(1, 20) = 11.7$. There were no other significant effects on error rates ($Fs < 1.7$). Correct positive responses to the yes–no questions averaged 1,747 ms (12% errors) and correct negative responses, 1,810 ms (23% errors).

Table 10
Results of Experiment 5: Response Times (in Milliseconds) and Error Rates on Test Words

	Test word			
	Critical noun		Location	
Text version	RT	% error	RT	% error
Critical noun				
With main character	1,078	13	1,148	27
Behind main character	1,144	11	1,203	24
	Filler test word			
Positive filler	1,344[a]		20[b]	
Negative filler	1,199[a]		14[b]	

Note. RT = response time.
[a] RT. [b] Percentage error.

Discussion

The Glenberg et al. (1987) and the Morrow et al. (1987; 1989) experiments fail to provide convincing evidence that real-life situation models are used automatically during comprehension because the results of both sets of experiments have interpretations that are consistent with the minimalist hypothesis. Of course, when readers have special goals or strategies, they can construct representations of quite complicated situations. Perrig and Kintsch's (1985) subjects were able to construct representations of the layout of a town from a text they read four times. Johnson-Laird's (1980) subjects, knowing they would have to draw a picture, were able to construct the relative positions of three objects from a textual description. As Johnson-Laird pointed out, constructing the information needed for a complete situation model can require considerable effort (see also Glenberg & Langston, 1992). If a passage in a story describes a complex scene with many interrelated objects, then a reader "would probably form only a rather vague idea of the actual spatial layout" (Johnson-Laird, 1980, p. 103). However, if the reader's goal were to answer a question about the relative location of a specific object, then the reader could use appropriate strategies and sufficient time during reading to construct the answer to the question. These strategies are clearly available to readers; the inferences required to represent lifelike information can be constructed. However, there is no empirical evidence to conclusively show that the inferences are constructed during reading by automatic processes.

An important conclusion to be drawn about constructionist inferences is that they contrast sharply with inferences that establish local coherence and inferences that make use of well-known information. Although these latter inferences (about propositional connections, reference, and well-known semantic relations) can be demonstrated easily in the prototypical laboratory experiment, there are no equivalently convincing demonstrations of the automatic encoding of real-life situational inferences.

General Discussion

It is widely believed that readers automatically construct inferences to build a relatively complete mental model of the situation described by a text (Glenberg et al., 1987; Johnson-Laird, 1980; Morrow et al., 1987; Rumelhart, 1975; Trabasso & van den Broek, 1985; van Dijk & Kintsch, 1983). However, our conclusion is that readers do not automatically encode the inferences that would make up such a model. We base this conclusion on several points:

1. The empirical evidence that has been put forward to demonstrate the automatic encoding of a life-like situation model can be explained by the minimalist hypothesis.
2. Elaborative inferences that should be part of a lifelike situation model, for example, instrumental inferences, are not explicitly and automatically encoded.
3. Global inferences to connect widely separated parts of a story are not automatically encoded.

A wide range of data has been shown to be consistent with the minimalist hypothesis. For local inferences based on information in working memory, the minimalist claim is that they will be encoded because they are quickly and easily available. This claim has been verified for the inferences that connect propositions through argument repetition and anaphora. Minimalist tests of other kinds of local inferences, for example, the minimalist causal inferences proposed by van den Broek (1990) and Fletcher (Fletcher & van den Broek, 1989), await further research. For inferences based on general knowledge, again the minimalist claim is supported. Inferences about the instruments taken by verbs, about the contextually relevant meanings of words, and about the prototypical members of categories are encoded if the information on which they are based is easily available during reading.

In contrast, there is little data to support the constructionist position. Experiments 1 through 4 showed that causal global inferences are not automatically encoded during reading. Inferences that might be assumed to be encoded under all circumstances, such as instrumental inferences, are not. Experiments that have been cited as verifying constructionist situation models (e.g., Morrow et al., 1987, 1989) have alternative interpretations (Alba & Hasher, 1983), and Experiment 5 demonstrated the validity of one such alternative interpretation.

Besides being consistent with data, the minimalist hypothesis has several advantages over the constructionist hypothesis. The minimalist hypothesis is falsifiable in that it has clearly testable predictions: An inference is not constructed unless it is necessary to establish local coherence or it is supported by well-known, easily available information. These predictions provide a direct focus for empirical tests of the hypothesis. In contrast, constructionists have rarely provided an account of exactly which inferences should be encoded and which should not be.

The minimalist hypothesis is also a vital alternative hypothesis that more elaborative theories will have to take into account before they can become serious theories of text processing. So long as there is no convincing body of empirical evidence to support the constructionist view of automatic encoding processes, then the minimalist hypothesis remains viable. By providing an alternative to an otherwise widely accepted view, the minimalist hypothesis can lead to potentially fruitful investigations of inference processes.

It should be stressed, as it has been throughout this article, that the controversy between the constructionist and minimalist views is about the inferences that a reader encodes automatically, in the absence of specific goals and strategies. Neither minimalist nor constructionist theories propose models of how strategic, goal-specific inference processing is carried out, and this issue remains on the agenda for future research.

Two strategies for research are suggested by the minimalist hypothesis. One strategy is broadly exploratory and based on the hypothesis that the kinds of information that support inferences are those that are easily available. Experiments are designed to test for the inferences that might be supported by a variety of different kinds of knowledge. Another general issue is the organization of the local information in a text, and so experiments are designed to investigate the relative availabilities of the different entities evoked by a text. Overall, the strategy is to hold to the minimalist hypothesis while searching as widely as possible for evidence to force its rejection.

The current empirical situation is that there is no conclusive evidence to support the constructionist hypotheses, and therefore no reason to reject the minimalist hypothesis. An exploratory strategy is one way to face the challenges imposed by this situation—many different kinds of inferences can be examined, and the wider the range, the more stringent will be the test of the minimalist claims.

The second strategy for research is to construct explicit models of minimalist processes and then test the models empirically. This is the strategy adopted by Kintsch (1988; Kintsch, Welsch, Schmalhofer, & Zimny, 1990; Kintsch & Welsch, 1991) and by van den Broek (1990) and Fletcher (Fletcher & van den Broek, 1989). Van den Broek and Fletcher's model for local coherence was discussed earlier in this article. It provides a definition of local coherence in terms of causality and assumes that if locally available propositions (those in short-term memory at the same time) provide adequate causal relations for each other, then no other, more global causal relations are constructed. Thus, it is an example of a minimalist processing model.

Kintsch (1988) has proposed a construction–integration model whereby the concepts stated in a text and information from general knowledge associated with the concepts all interact to produce an encoded representation of the text (see also Ratcliff & McKoon, 1988). The construction–integration process can both change the relations among propositions that are explicitly stated in a text and add propositions to the representation. The construction–integration process is explicitly defined: Integration of long-term memory information and text information takes place through a repeated recycling of activation so that information associated only weakly to a relatively small portion of a text is further weakened, whereas information associated more strongly to multiple concepts in the text is strengthened. This process can change the organization of explicitly stated propositions because long-term memory associations can strengthen connections between propositions that would otherwise be only weakly connected. In this sense, the integration process represents the text as a situation model in a way that is consistent with minimalist rather than constructionist claims.

The construction–integration process can also add inferences to the text representation. For example, if a text contains the word *mint*, strong associations to both meanings are immediately activated, providing support for potential inferences. If other concepts in the text are associated with one of the meanings and not the other, then activation of these associations leads to an increase in activation of associations to the contextually appropriate meaning of *mint* and a decrease in activation of associations to the contextually inappropriate meaning. This process results in the encoding of propositions (inferences) about the contextually appropriate meaning and implements the same claim as is made by the minimalist position, that information is added to a text representation to the extent that it is supported by easily available information.

Kintsch's (1988) model goes beyond the minimalist position in that it allows for the encoding of inferences that are based on information not immediately available. In the review of the literature presented in this article, the one finding inconsistent with the minimalist position was the finding that inferences about predictable events are encoded—not explicitly encoded as would be predicted by constructionist theories, but encoded to some degree. Kintsch showed how a minimalist processing model can account for this result. Inferences from long-term memory that are only weakly associated to individual concepts in a text can become more strongly activated over time if they are associated to several different concepts in the text. For the sentence "The townspeople were amazed to find that all the buildings had collapsed except the mint," the concept *earthquake* is not strongly associated with any individual concept in the sentence, and so it is not immediately available to support inferences. However, as the integration process proceeds, the recycling of activation from multiple sources may make this concept available as an inference to be added to the text representation.

Our goal in establishing the minimalist hypothesis is to stimulate research designed to find the principles by which inferences are generated. These principles might be defined across a number of kinds of inferences or within models of text processing. Either way, we believe that there will be two outcomes. First, both the minimalist and the constructionist hypotheses will be modified away from their all-or-none positions toward a more graded view of inference processing. This has already happened with inferences about predictable events. It cannot be said either that these inferences are completely and explicitly encoded or that they are not encoded at all. Instead, they are encoded to some degree, and finding evidence for them depends on finding the appropriate retrieval environment.

Second, the class of minimal inferences will be expanded. Currently, minimal inferences have been shown to include those that are supported by well-known semantic associations and well-known category membership relations. Expansion to include inferences that are supported by knowledge of the argument-taking properties of verbs has been tentatively suggested (Boland et al., 1990; Hudson, Tanenhaus, & Dell, 1987; McKoon & Ratcliff, 1989c; Tanenhaus et al., 1989). For example, in the sentence "he cleared the papers off" an argument of the verb (the place the papers were removed from) is missing. McKoon and Ratcliff (1989c) provided data indicating that readers inferred the missing argument from mention in a preceding sentence. Minimal inferences are also currently said to include those that are based on local information in a text. This notion too can be tentatively expanded toward a less simplistic view. Instead of regarding local information as a relatively undifferentiated list of propositions, we can borrow the discourse models used in artificial intelligence (Grosz et al., 1983; Sidner, 1983a, 1983b; Webber, 1983). According to these models, the information in a text is represented as the set of entities evoked by the text and the relations among them. Each entity in the model is assumed to have some degree of accessibility, which is determined by the syntactic, semantic, and pragmatic environment in which it is linguistically expressed. The varying degrees of accessibility should be reflected in the processes that construct local inferences. Initial evidence that inference processes are, in fact, affected in this way has been found in several studies (Greene et al., 1992; Hudson et al., 1987; McKoon, Ward, Ratcliff, & Sproat, in press; Ward et al., 1991). For example, comprehension of a pronoun is facilitated if the pronoun refers to an entity that is topical in its text, and if the referent entity was first mentioned in a salient syntactic position (Ward et al., 1991).

As the class of minimal inferences expands, the sharp contrast between the minimalist and constructionist positions may be redefined. For example, inferences previously thought to be constructed because they were necessary to a situation model may instead be understood to be based on easily available information and so become incorporated into a minimalist representation. At the same time, it may become clear which inferences cannot be constructed automatically, and for these inferences, models of strategic, goal-based generation processes will be required.

It is very important not to misunderstand the goal of the minimal inference position. It is easy to see it as a rejection of all goal-based, purposeful inference processing because this article is focused on minimal inferences. This is not the case. The aim is to try to separate the inferences and relations that are automatically and rapidly produced from those that are the result of slower, goal-based strategic processes. From such a separation, we can begin to understand the characteristics of the database provided in the first few hundred milliseconds of processing. Information about this database can then be used to tell us what information strategic processes have to work with (and therefore which strategic inferences will be difficult and which easy) and perhaps even identify strategic inferences that the processing system cannot avoid constructing.

References

Alba, J. W., Alexander, S., Hasher, L., & Caniglia, K. (1981). The role of context in the encoding of information. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 283–292.

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93,* 203–231.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* New York: Holt, Rinehart & Winston.

Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles—A problem of polysemy. *Cognitive Psychology, 7,* 167–180.

Anderson, R. C., Pichert, J. W., Goetz, E., Schallert, D., Stevens, K., & Trollip, S. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior, 15,* 667–679.

Baillet, S. D., & Keenan, J. M. (1986). The role of encoding and retrieval processes in the recall of text. *Discourse Processes, 9,* 247–268.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82–93.

Black, J. B., & Bower, G. H. (1980). Story understanding as problem solving. *Poetics, 9,* 223–250.

Bloom, C. P., Fletcher, C. R., van den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An online assessment of causal reasoning during comprehension. *Memory & Cognition, 18,* 65–71.

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language, 29,* 413–432.

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

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

Bransford, J. D., & 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.

Chang, F. R. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory & Cognition, 8,* 58–64.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition, 7,* 35–41.

Corbett, A. T. (1984). Pronominal adjectives and the disambiguation of anaphoric nouns. *Journal of Verbal Learning and Verbal Behavior, 17,* 683–695.

Corbett, A. T., & Chang, F. R. (1983). Pronoun disambiguation: Accessing potential antecedents. *Memory & Cognition, 11,* 283–294.

Corbett, A. T., & Dosher, B. A. (1978). Instrument inferences in sentence encoding. *Journal of Verbal Learning and Verbal Behavior, 17,* 479–491.

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

Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior, 22,* 121–132.

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

Dosher, B. A., & Corbett, A. T. (1982). Instrument inferences and verb schemata. *Memory & Cognition, 10,* 531–539.

Duffy, S. A. (1986). Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 208–219.

Ehrlich, S. F., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22,* 75–87.

Flagg, P. W. (1976). Semantic integration in sentence memory? *Journal of Verbal Learning and Verbal Behavior, 15,* 491–504.

Flagg, P. W., & Reynolds, A. G. (1977). Modality of presentation and blocking in sentence recognition memory. *Memory & Cognition, 5,* 111–115.

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

Fletcher, C. R., Hummel, J. E., & Marsolek, C. J. (1990). Causality and the allocation of attention during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 233–240.

Fletcher, C. R., & van den Broek, P. (1989). *A model of narrative comprehension and recall.* Manuscript submitted for publication.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33,* 465–495.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research, 18,* 51–60.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32,* 99–156.

Gerrig, R. J. (1986). Process models and pragmatics. In N. E. Sharkey (Ed.), *Advances in cognitive science.* Chichester, England: Ellis Horwood.

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

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

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

Graesser, A. C. (1981). *Prose comprehension beyond the word.* New York: Springer.

Graesser, A. C., Robertson, S. P., & Anderson, P. A. (1981). Incorporating inferences in narrative representations: A study of how and why. *Cognitive Psychology, 13,* 1–26.

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory,* and *Cognition, 18,* 266–283.

Greenspan, S. L. (1986). Semantic flexibility and referential specificity of concrete nouns. *Journal of Memory and Language, 25,* 539–557.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. *Proceedings of the 21st Annual Meeting of the Association of Computational Linguistics.* Association of Computational Linguistics.

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13,* 512–521.

Hudson, S. B., Tanenhaus, M. K., & Dell, G. S. (1987). The effect of the discourse center on the local coherence of a discourse. *Proceedings of the 8th Annual Cognitive Science Meetings.*

Johnson, M. K., Bransford, J. D., & Solomon, S. K. (1973). Memory of tacit implications of sentences. *Journal of Experimental Psychology, 98,* 203–205.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4,* 71–115.

Katz, S., & Gruenewald, P. (1974). The abstraction of linguistic ideas in "meaningless" sentences. *Memory & Cognition, 2,* 737–741.

Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior, 23,* 115–126.

Keenan, J. M., & Kintsch, W. (1974). The identification of explicitly and implicitly presented information. In W. Kintsch (Ed.), *The representation of meaning in memory* (pp. 153–176). Hillsdale, NJ: Erlbaum.

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

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95,* 163–182.

Kintsch, W., & Glass, G. (1974). Effects of propositional structure on text recall. In W. Kintsch (Ed.), *The representation of meaning in memory* (pp. 146–152). Hillsdale, NJ: Erlbaum.

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,* 257–274.

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

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

Kintsch, W., & Welsch, D. M. (1991). The construction–integration model: A framework for studying memory for text. In W. E. Hockley & S. Lewandowsky (Eds.), Relating theory and data: Essays on human memory (pp. 363–367). Hillsdale, NJ: Erlbaum.

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

Mandler, J. M. (1978). A code in the node: The use of a story schema in retrieval. *Discourse Processes, 1,* 14–35.

Mandler, J. M., & Johnson, N. J. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9,* 111–151.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10,* 181–187.

McKoon, G. (1977). Organization of information in text memory. *Journal of Verbal Learning and Verbal Behavior, 16,* 247–260.

McKoon, G., & Keenan, J. M. (1974). Response latencies to explicit and implicit statements as a function of the delay between reading and test. In W. Kintsch (Ed.), *The representation of meaning in memory* (pp. 166–176). Hillsdale, NJ: Erlbaum.

McKoon, G., & Ratcliff, R. (1980a). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior, 19,* 668–682.

McKoon, G., & Ratcliff, R. (1980b). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior, 19,* 369–386.

McKoon, G., & Ratcliff, R. (1981). The comprehension processes and memory structures involved in instrumental inference. *Journal of Verbal Learning and Verbal Behavior, 20,* 671–682.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 82–91.

McKoon, G., & Ratcliff, R. (1988). Contextually relevant aspects of meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 331–343.

McKoon, G., & Ratcliff, R. (1989a). Assessing the occurrence of elaborative inference with recognition: Compatibility checking vs. compound cue theory. *Journal of Memory and Language, 28,* 547–563.

McKoon, G., & Ratcliff, R. (1989b). Inferences about contextually defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1134–1146.

McKoon, G., & Ratcliff, R. (1989c, November). *Inferences based on lexical information about verbs.* Paper presented at the 30th Annual Meeting of the Psychonomic Society, Atlanta, GA.

McKoon, G., & Ratcliff, R. (1989d). Semantic association and elaborative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 326–338.

McKoon, G., & Ratcliff, R. (1989e). Textual inferences: Models and measures. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 403–421). Hillsdale, NJ: Erlbaum.

McKoon, G., & Ratcliff, R. (1990). Dimensions of inference. In A. Graesser & G. Bower (Eds.). *The psychology of learning and motivation* (Vol. 25, pp. 313–328). San Diego, CA: Academic Press.

McKoon, G., & Ratcliff, R., & Seifert, C. (1989). Making the connection: Generalized knowledge structures in story understanding. *Journal of Memory and Language, 28,* 711–734.

McKoon, G., Ward, G., Ratcliff, R., & Sproat, R. (in press). Morphosyntactic and pragmatic factors affecting accessibility of discourse entities. *Journal of Memory and Language.*

Moeser, S. D. (1976). Inferential reasoning in episodic memory. *Canadian Journal of Psychology, 31,* 41–70.

Morrow, D., Bower, G., & Greenspan, S. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28,* 292–312.

Morrow, D., Greenspan, S., & Bower, G. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26,* 165–187.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26,* 453–465.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18,* 5–20.

Omanson, R. C. (1982a). An analysis of narratives: Identifying central, supportive and distractive content. *Discourse Processes, 5,* 195–224.

Omanson, R. C. (1982b). The relation between centrality and story category variation. *Journal of Verbal Learning and Verbal Behavior, 21,* 326–337.

Paris, S., & Lindauer, B. K. (1976). The role of inference in children's comprehension and memory for sentences. *Cognitive Psychology, 8,* 217–227.

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

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inference: Lexical decision versus naming. *Journal of Memory and Language, 27,* 399–415.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior, 17,* 403–417.

Ratcliff, R., & McKoon, G. (1981a). Automatic and strategic priming in recognition. *Journal of Verbal Learning and Verbal Behavior, 20,* 204–215.

Ratcliff, R., & McKoon, G. (1981b). Does activation really spread? *Psychological Review, 88,* 454–462.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385–408.

Reitman, J. S., & Bower, G. H. (1973). Storage and later recognition of exemplars of concepts. *Cognitive Psychology, 4,* 194–206.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346–378.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science.* San Diego, CA: Academic Press.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. Laberge & J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 263–303). Hillsdale, NJ: 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.

Seifert, C. M., McKoon, G., Abelson, R. P., & Ratcliff, R. (1986). Memory connections between thematically similar episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 220–231.

Sidner, C. L. (1983a). Focusing in the comprehension of definite anaphora. In M. Brady & R. Berwick (Eds.), *Computational models of discourse.* Cambridge, MA: MIT Press.

Sidner, C. L. (1983b). Focusing and discourse. *Discourse Processes, 6,* 107–130.

Singer, M. (1978, August–September). *The role of explicit and implicit recall cues.* Paper presented at the 86th Annual Convention of the American Psychological Association, Toronto, Ontario, Canada.

Singer, M. (1979). Processes of inference during sentence encoding. *Memory & Cognition, 7,* 192–200.

Stein, N. L., & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing.* Hillsdale, NJ: Erlbaum.

Suh, S., & Trabasso, T. (1988, November). *Convergent evidence on inferences during comprehension of text.* Paper presented at the Annual Meeting of the Psychonomic Society, Chicago.

Swinney, D., & Osterhout, L. (1990). Inference generation during auditory language comprehension. In A. Graesser & G. Bower (Eds.), *The psychology of learning and motivation* (Vol. 25, pp. 17–33). San Diego, CA: Academic Press.

Tabossi, P. (1982). Sentential context and the interpretation of unambiguous words. *Quarterly Journal of Experimental Psychology, 34,* 79–90.

Tabossi, P., & Johnson-Laird, P. N. (1980). Linguistic context and the priming of semantic information. *Quarterly Journal of Experimental Psychology, 32,* 595–603.

Tanenhaus, M. K., Carlson, G. N., & Seidenberg, M. S. (1985). Do listeners compute linguistic representations? In D. Dowty, L. Kartunnen, & A. Zwicky (Eds.), *Natural language parsing.* Cambridge, England: Cambridge University Press.

Tanenhaus, M. K., Carlson, G. N., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4,* 211–234.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9,* 77–110.

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

Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language, 24,* 595–611.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24,* 612–630.

van den Broek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language, 27,* 1–22.

van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423–446). Hillsdale, NJ: Erlbaum.

van den Broek, P., & Trabasso, T. (1986). Causal networks versus goal hierarchies in summarizing text. *Discourse Processes, 9,* 1–15.

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

Ward, G., Sproat, R., & McKoon, G. A. (1991). Pragmatic analysis of so-called anaphoric islands. *Language, 67,* 439–474.

Webber, B. (1983). So what can we talk about now? In M. Brady & R. Berwick (Eds.), *Computational models of discourse.* Cambridge, MA: MIT Press.

Wilson, S. G., Rinck, M., McNamara, T. P., Bower, G. H., & Morrow, D. G. (1992). *Mental models and narrative comprehension: Some qualifications.* Unpublished manuscript.

Received September 10, 1990
Revision received March 20, 1991
Accepted June 12, 1991 ■