Journal of Experimental Psychology:
Learning, Memory, and Cognition
1989, Vol. 15, No. 2, 326–338

0278-7393/89/$00.75

Semantic Associations and Elaborative Inference

Gail McKoon and Roger Ratcliff
Northwestern University

In this article, a theoretical framework is proposed for the inference processes that occur during reading. According to the framework, inferences can vary in the degree to which they are encoded. This notion is supported by three experiments in this article that show that degree of encoding can depend on the amount of semantic-associative information available to support the inference processes. In the experiments, test words that express possible inferences from texts are presented for recognition. When testing is delayed, with other texts and test items intervening between a text and its test word, performance depends on the amount of semantic-associative information in the text. If the inferences represented by the test words are not supported by semantic associates in the text, they appear to be only minimally encoded (replicating McKoon & Ratcliff, 1986), but if they are supported by semantic associates, they are strongly encoded. With immediate testing, only 250 ms after the text, performance is shown to depend on semantic-associative information, not on textual information. This suggests that it is the fast availability of semantic information that allows it to support inference processes.

Previous investigations of inference and reading have usually centered on the question of whether some specific kind of information is inferred during reading. In this article, we propose a new, more general, framework within which to think about inference processes. The framework suggests a range of new questions about inference, and we show how these questions can be addressed empirically.

The first important proposed idea is that the question of whether an inference inferred during reading should be replaced by the question of what variables govern inference processes. Some variables may govern whether an inference is made at all; for example, elaborative inferences might occur only with specific goals of the reader. Other variables may determine the strength with which an inference is made, so that different kinds of inferences can be compared only in terms of relative degree of encoding and not in terms of one kind of inference being generated during reading and the other not. Still other variables may control the conditions under which evidence for an inference appears; for example, under some retrieval conditions, it may look as though an inference was encoded during reading, and under other conditions, it may look as though the inference was not encoded.

Consideration of such a range of variables leads to a second proposal, a framework in which an inference is not necessarily encoded in an all-or-none fashion but, instead, can be encoded partially. In this framework, an inference can be minimal, representing some set of features or propositions that does not completely instantiate the inference. Different kinds of inferences can be encoded with different amounts or strengths of information. If the strength of encoding is relatively high, then effects of the inference should appear under a variety of retrieval conditions. But if the strength is low, then effects of the inference may appear only under optimal retrieval conditions. After examining a range of specific retrieval conditions, it might be that the strength of some particular kind of inference would be considered high enough to describe it as explicitly encoded or that the strength of some other kind of inference would be considered so low that it could only be described as minimally encoded. More important, the minimal-inference framework suggests comparisons between different kinds of inferences under different kinds of retrieval conditions as a way of mapping the information included in the mental representation of a text.

To summarize, we propose that a useful framework for examining inference processes is one that stresses the continuum along which the degree of encoding of an inference can vary, from not encoded at all to minimally encoded to explicitly instantiated. This framework leads naturally to consideration of the variables that govern where on the continuum the encoding of any particular inference will fall.

In this article, we investigate the effects of one specific variable, semantic association. The claim is made that inference processes in general depend heavily on support from associative semantic information: If construction of an inference can make use of well-known semantic information, then the inference will be encoded with more strength or specificity than if construction of the inference must rely on information not so easily available. Support for this claim is provided by the experiment in the first part of this article. However, we want to go further than a simple demonstration that semantic-associative information affects inference processes; we want to begin to investigate the mechanisms that underly this effect, and we do that in the second part of the article.

Semantic-Associative Information

As a variable to affect inference processes, semantic association is an obvious candidate. The relations between highly

This research was supported by National Science Foundation (NSF) Grant BNS 85 10361 to Roger Ratcliff, National Institute of Health Grant HD 18812 to Gail McKoon, and NSF Grant BNS 85 16350 to Gail McKoon.

Correspondence concerning this article should be addressed to either Gail McKoon or Roger Ratcliff, Department of Psychology, Northwestern University, Evanston, Illinois 60208.

associated words are well-known to a reader, and they have been shown to be automatically and quickly accessible (cf. Neely, 1977; Meyer & Schvaneveldt, 1976), so it would not be surprising if inference processes made use of them.

There are at least two ways in which associative semantic information can interact with new information during reading. First, semantic associations may provide relations between two pieces of explicitly stated information, and these relations may contribute to the construction of inferred connections. If one sentence mentions the words *January, cold, colder*, and *water*, and the next sentence mentions the word *freeze*, then the associations between *freeze* and the prior words should help to relate the two sentences, and also *freeze* should be easier to understand in the context of the other words. The second way in which semantic information might support inferences during reading is to contribute to elaborative inferences. Elaborative inferences do not connect information explicitly stated in a text; rather they are inferences that add new, never stated information. Even if *freeze* is never mentioned explicitly, it might be more easily inferred because of the high association values between the words *cold, water*, and *freeze*.

Both of these functions of associative semantic information have been suggested by previous research. For explicitly stated information, the usual finding is that the higher the semantic association between the words in a context and the words to be related to that context, the easier the processing. Corbett (1984) and Garrod and Sanford (1977) have presented data that suggest that the difficulty of interpretation of a category name used an an anaphor depends on the typicality of the alternative referents, and Roth and Shoben (1983) have shown the dependence of relative typicality on context. More complicated processes are also affected by semantic information; Keenan, Baillet, and Brown (1984) have found that causal relations between sentences are easier to process when the sentences are more highly associated.

For elaborative inferences, there is little data about the effects of semantic association. One previous study does suggest that inferences about the instruments of verbs depend on the degree of association between the instrument and the verb (McKoon & Ratcliff, 1981). However, the instrument of the verb was explicitly presented, so that all the inference processes had to do was to connect an instrument stated at one point in a text to the appropriate verb stated at a later point in the text. The purpose of Experiment 1 in this article is to provide evidence that semantic association affects inference processes even when the to-be-inferred information is never stated in the text.

Experiment 1

The manipulation of semantic association used in the experiment is best described with reference to the examples of material shown in Table 1. The first sentence is a *predicting* sentence because it allows an elaborative inference about, or predicts, the target test word *dead*. The second sentence is a *control* sentence because it does not predict the target word *dead*, even though it contains many of the same words as the predicting sentence. Neither of these sentences contains any words that would be individually strongly associated to the target word. At most, there might be weak associations between the target and such words as *fell* and *14th story*. Therefore this pair of sentences is labeled as having *weak associations* to the target word.

Table 1
Examples of Materials for Experiment 1

Condition	Sentence	Target test word
	Weak associations	
Predicting	The director and cameraman were ready to shoot close-ups when suddenly the actress fell from the 14th story.	Dead
Control	Suddenly the director fell upon the cameraman, demanding that he get a close-up of the actress on the 14th story.	Dead
	Strong associations	
Predicting	The housewife was learning to be a seamstress and needed practice so she got out the skirt she was making and threaded her needle.	Sew
Control	The housewife was a careless seamstress, and when she dropped an unthreaded needle on the floor, she didn't find it until she stepped on it days later.	Sew

In contrast, the sentences of the pair labeled *strong associations* contain words that are individually strongly associated to the target word *sew*; these include *seamstress, thread*, and *needle*. It should be stressed that both the predicting and the control sentences contain these words so that intraword associations are held constant, but only the predicting sentence would lead to an inference represented by the target *sew*.

The hypothesis for Experiment 1 was that there would be more information that matched the target word included in the encoded representation of a predicting sentence if the sentence contained strong associates to the target than if it did not. Strong semantic associates could support the inference in several ways: The amount of information encoded for the inference could be greater or more specific, or the strong associates could make the inference more probable. These different interpretations will be examined in the final discussion section of the article, after presentation of the experiments.

Methodology and Background for Experiment 1

Elaborative inferences have been studied frequently, with the result that there is some understanding of different methods of investigation. For example, it is generally agreed that on-line methods by which a target test word is presented immediately following a text cannot be used to demonstrate elaborative inferences. This is because there is no way of separating forward, elaborative mechanisms from backward, context-checking mechanisms that occur at the time the test item is presented (cf. Forster, 1981). Similarly, measures of recall may be affected by information encoded when a text was read or by information constructed at the time of the recall test (Corbett & Dosher, 1978; Singer, 1978). To separate encoded information from information constructed by retrieval processes, McKoon and Ratcliff (1980a, 1980b, 1988c; Ratcliff & McKoon, 1978) have used delayed, speeded item recognition. The rationale is that backward context checking is ruled out because the test is delayed by the presentation of other material and that construction at the time of test is ruled out because the retrieval process is fast and automatic and test items relevant to the inferences under investigation are presented with a low probability (McKoon & Ratcliff, 1988a; Ratcliff & McKoon, 1981).

Using speeded item recognition, McKoon and Ratcliff (1986, 1987) provided evidence that some elaborative inferences are encoded minimally at the time of reading a text. In the experiments, subjects read short texts in either predicting or control versions. There was a series of study-test trials (shown by example in Table 2), each with two unrelated texts to read in the study phase and a list of words for recognition in the test phase. Recognition test words were presented as prime–target pairs. The prime was either the neutral word *ready* or a word from one of the studied texts. The target for the materials of interest was a word representing the event predicted by the predicting sentence; the correct response to this word was *no* because it had not appeared in the studied sentence. McKoon and Ratcliff (1986, 1987) found that responses to the target were inhibited in the predicting condition relative to the control condition when the prime was from the studied text but that there was no significant inhibition when the prime was the word *ready*. These results were interpreted as evidence for a minimal inference that was formed during encoding of the text, where a minimal inference might be something like *something bad happened* for the text about the actress falling from the 14th story. The idea was that the inference was so minimal that the test word *dead* did not match it sufficiently to give inhibition when *dead* was presented alone (in the neutral condition) but that the test word did match sufficiently when it was presented in combination with a prime from the sentence. It was this finding, that elaborative inferences could be encoded minimally and appear only under certain retrieval conditions, that led to the proposal of the framework emphasizing minimal and variable inference processes.

Table 2
Example of a Study-Test Trial

Press space bar to begin

The director and cameraman were ready to shoot close-ups when suddenly the actress fell from the 14th story.

The housewife was a careless seamstress, and when she dropped an unthreaded needle on the floor, she didn't find it until she stepped on it days later.

TEST TEST TEST

++++++++

ready

table

++++++++

housewife

sew (target test word primed by word from sentence)

++++++++

floor

stepped

++++++++

ready

dead (target test word primed by neutral word)

++++++++

ready

cameraman

.

.

.

.

.

.

Experiment 1 used the same procedure as McKoon and Ratcliff (1986). There were two sets of materials, those with weak associations between words of the text and the target test word and those with strong associations. Some of the weak-association materials were taken from McKoon and Ratcliff (1986), some of them were modifications of texts from McKoon and Ratcliff (1986) with associates of the target words removed from the text, and some were new. The strong-association materials were written to include in each sentence as many associates of the test word as possible.

The design of the experiment followed that by McKoon and Ratcliff (1986). Subjects read either the predicting or the control version of a sentence, and the target test word was primed either by the neutral word *ready* or by a word from the studied sentence.

For the weak-associations texts, the hypothesis was that the inference represented by the target test word would be minimally encoded in the mental representation of the predicting sentence. This means that the match between the target and the predicting text representation in memory would be poor, unless the target was combined with a prime word from the sentence. With the neutral prime, performance on the target should be about the same for the predicting and control texts. But with the prime from the sentence, the combination of prime and target should match the predicting text well enough to give inhibition for the predicting texts relative to the control texts. Thus, with the weak-association materials, evidence for the encoding of the inference should appear only when retrieval conditions combine the target with the prime from the sentence.

For the strong-association texts, the hypothesis was that the words in the text that are strongly associated to the target would support the inference process for the predicting sentence. The match between the encoded representation of the predicting sentence and the test word should be good, whatever the prime. Thus, there should be more inhibition with the predicting than the control texts, both with the neutral prime and with the prime from the sentence. In other words, evidence for the encoding of the inference should appear in both priming conditions.

Method

Subjects. There were two groups of subjects, 36 per group, one group for the strong-association materials and one for the weak-association materials. The subjects participated either for credit in an introductory psychology course or pay of $5 for the 1-hr session.

Materials. There were 28 pairs of weak-association sentences and 28 pairs of strong-association sentences. The first sentence of each pair predicted the event represented by the test word. The second sentence used as many as possible of the same words as the first (including all words highly associated to the test word) but did not predict the event represented by the test word. For each pair, there were four test words; one was the predicted event, another was the main character, and two were other words, each used in both sentences. The sentences varied from one to three lines in length, as presented on the CRT screen, and varied from 10 to 27 words in length; the ranges were not different across experimental conditions.

There were also 28 filler paragraphs, each with four positive test words and two negative test words. Half of the filler paragraphs were two lines long as presented on the CRT screen, and half were three lines long.

Procedure. Subjects were tested in one 50-min session each. Presentation of all materials and collection of data were controlled by a real-time microcomputer system.

To ensure that subjects responded quickly enough so that slow, strategic processes could be ruled out, a deadline procedure was used (as in McKoon & Ratcliff, 1986). Subjects were given practice at responding to a deadline with a lexical-decision task. Each test item in the lexical-decision task began with a row of +s displayed for 500 ms. Then the +s were replaced by a prime, either *ready* or some other word, displayed for 200 ms. When the prime disappeared, a test letter string appeared on the next line. After the letter string had been displayed for 250 ms, a row of asterisks was presented below it. Subjects were instructed to respond exactly 300 ms after the asterisks appeared, "yes" with the *?*/key if the string was a word, "no" with the *Z* key if the string was not a word. After the response, the response time was displayed for the subject for 750 ms. Then, after a 500-ms pause, the next test item began. After every 10th test item, there was an instruction to press the space bar when ready to begin the next 10 items. There were 170 test items altogether. Subjects reported feeling comfortable with the deadline procedure after this much practice.

After the lexical-decision practice, subjects began the study-test procedure. Each subject was presented with 4 practice trials, followed by 28 experimental trials. On each trial, there were two texts to study and six test items.

A trial began with an instruction to the subject, printed on the CRT screen, to press the space bar to initiate the trial. Then the two study texts were presented, one at a time. Two-sentence filler paragraphs were presented for 5 s, and three-sentence fillers were presented for 6 s. One-line experimental sentences were presented for 4 s, two-line experimental sentences were presented for 6 s, and three-line experimental sentences were presented for 8 s. After each study text, there was a 2-s pause before the next text was presented. After the pause following the second text, a warning signal that the test list was about to begin was shown for 1.5 s.

In the test list, each test item began with a row of +s, shown for 500 ms. Next, the +s were replaced by the prime word, shown for 200 ms. Then the prime disappeared, and the test word was displayed on the line below. A row of asterisks appeared underneath the test word 350 ms after it was presented. Subjects were instructed, as in the lexical-decision practice, to give their response exactly 300 ms after the asterisks, responding yes if the test word had appeared in a studied sentence, and no if it had not. If the subject made an error, the word *error* was presented for 750 ms. Response time was displayed just as in the lexical-decision practice, and then the next test item began after 500 ms. After the sixth test item, the instruction to press the space bar to begin the next trial was displayed. Subjects were told to attend to the primes because they might facilitate responding to the test words.

On each experimental trial, one of the texts to be studied was a filler item and one an experimental item. One of the six test words was the word predicted by the predicting version of the experimental sentence in the study list. It was primed either by the main character in the sentence or by the word *ready.* Another of the six items consisted of either two words from the experimental sentence, one as prime and one as test word (probability = .75), or one word from the experimental sentence, primed by the word *ready* (probability = .25). For the filler text, there were four test items. For two of them, either the prime and test word were both from the text (probability = .75), or the prime was the word *ready,* and the test word was from the text (probability = .25). For the other two, either the prime was from the text and the test word was not from any text (probability = .75) or the prime was the word *ready* and the test word was not from any text (probability = .25). Thus, of the six test items, three required a positive response and three a negative response. No word was repeated in the test list (except for the neutral prime, *ready*).

A different random order of presentation of study and test items was used for every second subject. There were two restrictions on a test list: The test word that was predicted by the predicting version of the experimental sentence could not appear as the first test word, and other words from the experimental sentence had to appear later in the test list than the predicted word.

Design. There were two separate experiments, each with 36 subjects, one for the strong-association materials and one for the weak-association materials. In each experiment, there were four experimental conditions formed by crossing sentence type (predicting or control) and prime type (word from the sentence or *ready*). These conditions were combined with groups of subjects (9 per group) and sets of material (seven per set) in a Latin-square design.

Results

Means were calculated for each subject and each item in each condition, and means of these means are shown in Table 3. Keep in mind that response times are measured from the point at which the signal to respond is displayed, which was 350 ms after the display of the test word. There were no significant differences in response times across any conditions.

Table 3
Results of Experiment 1

	Prime			
	Text word		Neutral word	
Association	% error	Correct RT (ms)	% error	Correct RT (ms)
Weak				
Predicting sentence	48	373	34	352
Control sentence	21	362	23	339
Strong				
Predicting sentence	64	388	59	357
Control sentence	40	382	40	340

Note. RT = response time.

Thus, differences among experimental conditions appear in error rates. For all significant *F*s, $p < .05$ unless otherwise noted.

For the materials with weak associations, the results matched the predictions, so it appears that inferences about predicted events were minimally encoded in the mental representations of the predicting sentences. The inferences could not have been strongly encoded because the target words presented alone did not match the mental representation of the predicting sentences much better than they matched the representation of the control sentences. But the inferences must have been encoded to some extent because when the targets were combined with the primes from the sentences, they matched the representations better for the predicting than for the control sentences.

The data show this pattern in that a prime from a sentence led to a greater difference in amount of inhibition between the predicting and control sentences than did the neutral prime. This interaction was significant by analysis of variance (ANOVA), $F(1, 35) = 8.76$, with subjects as the random variable, and $F(1, 27) = 5.95$, with items as the random variable. The main effect of prime word was also significant, $F(1, 35) = 41.2$, and $F(1, 27) = 21.1$, and the main effect of sentence type (predicting vs. control) approached significance in the subjects analysis, $F(1, 35) = 3.25$, $p < .082$, and was significant in the items analysis, $F(1, 27) = 5.06$.

These results essentially replicate those of McKoon and Ratcliff (1986). The one difference is in the neutral priming condition. In this condition, McKoon and Ratcliff found a 9% difference in error rate between the predicting- and control-sentence conditions, which was not significant. The difference in the current experiment is 11%, about the same size, but significant by post hoc test with subjects as the random variable, $F(1, 35) = 7.8$, and approaching significance with items as the random variable, $F(1, 27) = 3.4$, $p < .10$. McKoon and Ratcliff (1986) interpreted the difference as suggesting that for some target words and some subjects, the target word matched the encoded text representation strongly enough that inhibition occurred even without a prime from the text. The current result gives support to that interpretation.

In sum, the conclusion for target words weakly associated to their texts is clear: The difference in the amount of inhibition for predicting versus control texts is larger with the word from the text as prime than with the neutral prime. This finding is demonstrated both by the results reported here and by the previous results reported by McKoon and Ratcliff (1986).

In contrast, for target words strongly associated to their texts, the amount of inhibition on the target word did not depend on whether the prime was neutral or from the sentence. Predicting sentences led to about the same amount of inhibition in both priming conditions. Apparently, as hypothesized, the words in the predicting sentence that were strongly associated to the target word supported the inference, and so the target word matched the mental representation of the predicting sentence strongly enough to give as much inhibition when it was presented by itself, in the neutral priming condition, as when it was presented with the prime from the sentence. An ANOVA showed that the difference due to sentence type, predicting versus control, was significant, $F(1, 35) = 47.3$, and $F(1, 27) = 40.8$, and that this difference did not depend on priming conditions, $Fs < 1.0$ for the interaction of priming condition and sentence type.

The results of Experiment 1 are exactly as predicted from the idea that semantic association can increase the strength of encoding of an inference. However, this experiment demonstrates this effect with only one set of materials and one kind of inference. To give greater confidence in the effect, we included a similar manipulation in another experiment (McKoon & Ratcliff, 1988c). The materials in this experiment were also texts that contained words highly associated to their test words, but the association was one of category membership. For example, a text might discuss a granddaughter milking an animal on a farm, and the test word would be *cow*. Using the same procedure as in the current experiment, the amount of inhibition on the test word was greater with the predicting than with the control texts, and the difference did not depend on priming condition. Thus, the experiment gives further support to the notion that the strength of encoding of an inference can be increased by semantic association.

In addition to the results of interest in Experiment 1, data were also tabulated for the filler test items, and these data show that performance was about the same for the two groups. For the group of subjects in the experiment with the weak-association materials, mean response time for positive test items primed by a word from a studied text was 306 ms (13% errors) and primed by *ready*, 322 ms (16% errors). For negative test items, the corresponding numbers were 339 ms (14% errors) and 324 ms (17% errors). For the group of subjects with the strong-association materials, results for positive test items with a prime from a studied text were 305 ms (17% errors), and with *ready* as a prime, 316 ms (20% errors). For negative responses, the means were 330 ms (15% errors) and 328 ms (14% errors).

Discussion

The main result of Experiment 1 is the contrast between the strong- and weak-association materials. In the strong-association case, a predicted test word matches the memory representation of its predicting sentence to such an extent that responses for the test word are inhibited with the test word alone (in the neutral priming condition) about as much as with the prime from the text. In the weak-association case, the match is not as strong with the test word alone as with the prime from the text. In other words, evidence for the encoding of the weak-association inferences depends on retrieval conditions, but evidence for the strong-association inferences does not.

We would like to interpret these results in terms of semantic associations supporting inference processes, but there is one possible alternative explanation. It may be that the materials differ, not in terms of weak versus strong semantic association, but in terms of predictability of the events represented by the test words. It may be that the events represented by the test words are more likely outcomes for the predicting sentences in the strong-association materials than in the weak-association materials.

If the two sets of materials differ in terms of predictability, then this difference should appear in measures other than the recognition error rates that were the focus of Experiment 1. We obtained two such measures. First, subjects were asked to rate (on a scale from 1 to 7) how predictable the predicted test word was, given the predicting sentence or (for different subjects) given the control sentence. The average ratings were only slightly higher for the strong-association materials than for weak-association materials (6.39 and 2.88, predicting and control, respectively, for the strong-association materials; and 5.89 and 2.68 for the weak-association materials). This is a small difference, given that subjects were probably not rating predictability alone but, instead, predictability plus semantic associations between sentence and predicted word.

The other way predictability was measured was to count the number of different responses given by subjects when asked to write "what happened next" continuations after predicting sentences. This task cannot reflect what inferences subjects generate when they are not asked to write continuations because writing a continuation forces the generation of an inference specific enough to write down. But large differences between the two sets of materials with respect to the continuation task still might be cause for concern. In fact, there was only a small difference between the two sets: For the strong-association predicting sentences, 92% of responses mentioned the predicted event, either explicitly or implicitly, and for the weak-association sentences, 86% mentioned the predicted event. Furthermore, this difference can be shown not to account for the difference obtained in the recognition error rates. Using the continuation responses, we picked the 10 weak-association items most like the strong-association items and tabulated their data from Experiment 1 separately. These data were almost identical to the data from all 28 weak-association items, with differences in error rates between the two sets of data no larger than 2%. Thus, the different results for the strong- and weak-association materials cannot be explained by differences in predictability, when predictability is measured by asking subjects to generate what happened next. Instead, we interpret the differences between the results for the strong- and weak-association materials in terms of the strength and specificity of inferred information. This explanation will be taken up in the Conclusions section of this article.

Experiment 2

According to our framework for understanding inference processes, the encoding of an inference can depend on many variables. Experiment 1 showed that one such variable is the degree of semantic association between the words in a text and the to-be-inferred information. When the degree of association is high, an inference is more strongly encoded than when it is low. Experiment 2 was designed to investigate why it is that degree of association should have such an effect. One possible reason is that semantic-associative information is available quickly. So Experiment 2 examined whether relations that depend on semantic associations are available at a point when relations without such associations are not yet available.

To control time so that availability could be measured, we used the procedure shown in Table 4. Subjects initiated the presentation of each sentence by pressing the space bar of a CRT keyboard. Then the words of the sentence were presented one at a time across the CRT screen, with each word appearing 250 ms after the preceding word. At the end of the sentence, all the words of the sentence were erased, and a test word appeared for recognition.

With this procedure, it is possible to control the time between presentation of the test word and presentation of words of the text that might lead to an inference involving the test word. For example, for the first sentence of Table 4, the relation between *hitting the cement* and *hurt* can be inferred only when *cement* is presented. So the time available to relate *cement* and *hurt* is only 250 ms (plus whatever portion of the response time can be used for further processing). If the relation is computed that quickly, then it will tend to make correct negative responses to *hurt* difficult and thus slow or inaccurate.

This on-line recognition procedure was used in Experiment 2 to compare the availability of relations that had strong support from semantic associations with the availability of relations that did not have such support. Examples of all the kinds of materials are shown in Table 5. For the first pair of sentences, predicting and nonpredicting control, the relation between the final word of the predicting sentence and the test word is not supported by semantic associations. For the second and third pairs of sentences, the relation is supported by semantic associations; *sit* is highly associated to *chair*. The control sentence for the second pair does not have any words related to the target and does not predict the target. In contrast, the control for the third pairs does have a semantic associate but does not predict the target. For this sentence,

Table 4
An Example Showing the Procedure of Experiment 2

Press space bar to begin

The diver prepared to do a double somersault into the pool; he jumped, spun, and hit the cement. HURT

Press space bar to begin

After shopping for hours, the grandmother headed for her favorite chair. SIT

Table 5
Examples of Materials for Experiment 2

Condition	Sentence	Target test word
	Weak associations	
Predicting	The diver prepared to do a double somersault into the pool; he jumped, spun, and hit the cement.	Hurt
Nonpredicting control	The diver prepared to do a double somersault into the pool; he jumped, spun, and hit the water.	Hurt
	Strong associations	
Predicting	After shopping for hours, the grandmother headed for her favorite chair.	Sit
Nonsemantic control	After shopping for hours, the grandmother headed for her favorite store.	Sit
	Strong associations	
Predicting	After shopping for hours, the grandmother headed for her favorite chair.	Sit
Nonpredicting control	After shopping for hours, the grandmother finally found the perfect chair.	Sit

the target is related to the final word of the sentence by semantic association, but it is not related to the meaning of the sentence as a whole; the grandmother is not likely to sit.

Predictions for these different kinds of materials can be made from the hypothesis that relations depending on semantic-associative information are available at a point when relations without semantic information are not available. For the first pair of sentences, the final word of the predicting sentence (*cement*) is not semantically related to the test word (*hurt*). If information to relate these two words is not available at the time of test (i.e., within 250 ms), then responses to *hurt* will not be inhibited, and there will be no differences in performance between the predicting and nonpredicting control sentences. For the second pair of sentences, the final word of the predicting sentence (*chair*) and the test word (*sit*) are semantically related, and this relation should be available in 250 ms. So responses should be inhibited with the predicting sentence relative to the nonsemantic control sentence. For the third pair of sentences, both the predicting and the nonpredicting control sentences end with a word (*chair*) that is semantically related to the test word (*sit*). If the semantic relation is available, and the relation between a sentence as a whole and the test word is not available, then responses should be inhibited for both sentences. In sum, relations that involve semantic association are predicted to affect responses, but relations that do not involve semantic association will not affect responses.

Two aspects of Experiment 2 should be stressed. One is that the task is recognition; the subjects must decide whether the test word explicitly stated in the sentence just read. Previous research has indicated that inferences do not necessarily affect immediate tests when the task is lexical decision (Glenberg, Meyer, & Lindem, 1987; Kintsch & Mross, 1985; McKoon, 1988), and it has been suggested that the inference processes involved with building a text representation do not affect the lexical-access processes needed for word identification (Kintsch & Mross, 1985). Recognition, it contrast with lexical decision, requires that subjects access the text representation and so should provide a measure of the availability of information contained in that representation.

The second point about Experiment 2 is that the procedure cannot distinguish forward, predicting inference mechanisms from backward mechanisms that begin only after the test word has been presented. It might seem that a backward mechanism would have to predict that inhibition should be found no matter what the time interval between sentence and test word is. But it might be that an immediate (250 ms) test is too quick for a complete representation of the text to have been constructed, so that the backward process has a poor representation to work with. Similarly, a forward process that operated quickly would predict inhibition for any time interval, whereas a forward process that was slow would predict no inhibition at immediate test. Thus, the results of Experiment 2 will not distinguish between forward and backward mechanisms; instead, the results will give a picture of availability of the information necessary to give inhibition, where that information may have been constructed by forward or backward mechanisms, or both.

Method

Subject. For the comparison of sentences with weak associations (predicting and nonpredicting control), there were 36 subjects. For the other two comparisons, one between the strong-association predicting and nonpredicting control and one between the strong-association predicting and nonsemantic control, there were 20 subjects each. All subjects participated in the experiment for credit in an introductory psychology course or for $5 for the 1-hr session. Subjects were randomly assigned to one of the three comparisons at the time they appeared in the lab.

Materials. There were 16 pairs of weak-association sentences. The sentences were chosen from those used by McKoon and Ratcliff (1986) and were modified so that the event represented by the test word could be predicted given the last word of the predicting sentence and not predicted without that word. There were also 16 triples of sentences with strong associations. The predicting sentence of the triple predicted the test word event, only at the last word of the sentence, and the last word was a high-semantic associate of the test word. The second sentence of the triple did not predict the test word but did use the same, semantically associated last word. The third sentence of the triple used as many as possible of the same words as the other two, but the last word was changed so that the test word was not predicted and so that the last word was not a semantic associate of the test word. All of these sentences were two lines long, as presented on a CRT screen, and varied in number of words from 12 to 21.

The sentences of the weak-association pairs differed from one another only in the final words of the sentences. All other words were held constant. The sentences of the strong-association triples were the same in that only the final word of a sentence was changed in order to vary semantic association. However, when semantic association was held constant (in the third pair of sentences in Table 5), the final word of the sentence had to be held constant. So, to vary predictability of the test word from the sentence as a whole, other words were changed. For example, in Table 5, *headed for her favorite chair* was changed to *finally found the perfect chair.* These sentences were kept as similar as possible, with only one to three words changing.

The test word for the experimental materials just described was always the event predicted by the predicting version of its sentences, and for the strong associates, the test word was associated to the final word of the sentence according to association norms (chosen from the list in McKoon & Ratcliff, 1979). The correct response for the test word was always negative, and the test word was always presented at the end of the sentence. Filler sentences were used for positive test words and for negative test words not associated to the sentences. There were 36 filler sentences: Ten were two lines long and had positive test words, and 26 were three lines long, 10 with negative test words and 16 with positive test words. For half of the filler sentences, a test word was presented toward the beginning of the final line, before the end of the sentence. For the other half, the test word was presented at the end of the sentence. Each filler sentence also had associated with it a true/false test sentence. These were designed to encourage subjects to read for comprehension, and they were clearly true or obviously false, half of each.

Procedure. All stimuli were presented on a CRT screen controlled by a real-time computer system. Subjects made their response by pressing keys on the CRT keyboard. ?/ for "yes" and Z for "no."

So that subjects would be able to respond quickly when given a test word following a sentence, they were given practice with a deadline procedure. A list of 100 lexical-decision test items was used. Each item was made up of a warning signal (a row of +s) displayed for 500 ms, a prime word displayed for 200 ms, and then a string of letters. The signal to respond was a row of asterisks presented 250 ms after the string, just below the string. The subjects were instructed to time their responses to exactly 300 ms after the asterisks appeared, pressing the /? key for *word* and the Z key for *nonword.* If the response was an error, the word *error* was displayed for 2,000 ms. Whether or not the response was correct, the response time was displayed for 750 ms. Then, after a 500-ms pause, the warning signal for the next item was presented. After each 10 items, an instruction was presented that told subjects to press the space bar to begin the next 10 items.

After this lexical-decision practice, subjects were given practice with 30 sentences followed by test words, and then the experiment proper began. Each sentence began with an instruction of the CRT screen to press the space bar on the keyboard. When the space bar was pressed, there was a 500-ms pause, and then the words of the sentence were displayed sequentially across the screen, each word appearing 250 ms after the previous one. When the words reached the end of a line, they continued on the next line below. After the last word of the sentence had been displayed for 250 ms, the entire sentence disappeared from the screen, and the test word was displayed in the position where the next word would have appeared, in capital letters. The signal to respond (the row of asterisks) was presented 300 ms after the test word, just below it. Response time and error feedback were given in the same way as in the lexical-decision practice. After feedback, the instruction to press the space bar for the next sentence appeared on the CRT screen.

After every sixth sentence, two test sentences were given for true/false verification. These sentences were presented all at once, not word by word, and subjects were told that only accuracy, not speed, was important.

Design. The three separate comparisons, each with a different group of subjects, were all designed in the same way. There were two experimental conditions, one in which the sentence was predicting and one in which it was one of the controls. These were combined with sets of sentences (eight per set) and groups of subjects in a Latin-square design. Order of presentation of sentences was random (different for each second subject), except that each six sentences contained four fillers.

Results

Means for speed and accuracy of responses to test words were calculated for each subject in each condition. The response times were measured from the onset of the response signal to the subject's keypress; they did not include the 300 ms from presentation of the test word to the response signal. We had intended the deadline procedure to force differences in experimental conditions into error-rate differences, by requiring subjects to keep their response times constant. After piloting several different variations on the deadline procedure, it became apparent that subjects found this too difficult in conjunction with the on-line reading task, probably because of the necessity of switching from reading at 250 ms per word to responding to a test word. However, we retained the deadline, as described in the Method section, to keep subjects responding quickly enough that we could be confident that they were not using time-consuming strategies for their responses. Thus, the data of main interest turn out to be response times. The data are shown in Table 6.

For the weak-association sentences, there was no significant inhibition for the responses following predicting sentences relative to the responses following control sentences. This result suggests that the relation between the predicted test word and the predicting sentence was not computed in the time available. Although mean responses were slower in the predicting condition, the difference did not reach significance even though we tested more subjects for this comparison (36 subjects) than for the comparisons with the strong-association sentences (20 subjects each). For correct responses, the mean responses times were not different statistically, $F(1, 35) = 2.7$,

Table 6
Results From Experiment 2

	Predicting		Nonpredicting control		Nonsemantic control	
Association	RT (ms)	% error	RT (ms)	% error	RT (ms)	% error
Weak	468	18	448	15	—	—
Strong	458	16	—	—	393	13
Strong	441	27	462	27	—	—

Note. RT = response time.

with subjects as the random variable, and $F(1, 15) = 2.2$, with items as the random variable. The error rates were also not different statistically, $F(1, 35) = 1.8$, and $F(1, 15) = 1.06$. Standard error of the response time means was 11 ms.

For the strong-association sentences, there were two comparisons. First, the predicting sentence was compared with the nonsemantic control, which did not predict the test word event and did not end with a word associated to the test word. With the predicting sentence, responses to the test words were slower and less accurate than with the control sentence. This difference shows inhibition for the test word when it is semantically related to the text and predicted by the text relative to when it is not semantically related or predicted.

Second, for another group of subjects, the predicting sentences were compared with the nonpredicting control sentences; both kinds of sentences ended with a word highly semantically associated to the test word, and both led to inhibition. Comparing against the nonsemantic control for the first group of subjects, responses for the second group of subjects for both the predicting and the nonpredicting sentences were slower (441 ms and 462 ms) than for the nonsemantic control (which was 393 ms), and both had much higher error rates (27% vs. the nonsemantic control at 13%). Comparing the predicting sentences for both groups of subjects against the nonsemantic control, response times were slower than the nonsemantic control for both groups (441 ms and 458 ms), and error rates were higher than the nonsemantic control for both groups (27% and 16%). The 27% error rate shows a particularly large amount of inhibition for the second group, perhaps because this group had somewhat higher error rates overall, as shown by the data on filler test items below. Combining the two groups of subjects, there appears to be inhibition in every condition in which the final word of the text was semantically associated with the test word. The only condition in which there appears to be no inhibition is the condition without a semantically associated final word, the nonsemantic control.

This interaction was shown significant by an ANOVA. For the two strong-association comparisons, the two groups of subjects were combined in a between-groups ANOVA with two within-subjects factor: comparison type (predicting vs. nonpredicting or predicting vs. nonsemantic) and sentence type (predicting or control). The interaction between these two factors was significant for response times with subjects as the random variable, $F(1, 38) = 3.8$, $p < .06$, and with items as the random variable, $F(1, 30) = 4.6$. For error rates, the comparison types (different groups of subjects) were significantly different, $F(1, 38) = 5.0$, with subjects, and $F(1, 30) = 10.7$, with items. All other Fs were less than 2.4. The standard error of the response time means was 14 ms.

To summarize the data, for the weak-association sentences, there was no significant inhibition for the predicting sentences relative to the control sentences. On the other hand, inhibition was obtained for the strong-association-predicting sentences and the strong-association-nonpredicting sentences relative to the strong-association-nonsemantic control sentences. In other words, significant inhibition was obtained when and only when the final word of the sentence was highly semantically related to the test word.

Overall, data from filler items show that the three groups of subjects were roughly comparable, although as mentioned earlier, the second group had somewhat higher error rates. For subjects with the weak-association sentences, the mean response time for positive filler test words was 404 ms (23% errors), and the mean for negative test words was 430 ms (16% errors). For subjects with the strong-association comparison of predicting and nonpredicting control sentences, the mean for positive filler test words was 404 ms (24% errors) and for negative test words, 396 ms (17% errors). For the subjects with the strong-association comparison of predicting and nonsemantic control sentences, the mean for positive test words was 379 ms (20% errors) and for negative test words, 401 ms (11% errors). The three groups of subjects also performed about the same on the true/false test sentences. In the same order, the first group made 26% errors on true sentences and 33% errors on false sentences. The second group made 26% and 32% errors, respectively, and the third group, 23% and 34%.

Discussion

The results of Experiment 2 show that the information relating two highly associated words, one at the end of a sentence and one a test word, is available to inhibit a recognition decision even when only 250 ms separates presentation of the two words. This immediate effect of semantic association is what would be expected from previous findings with lexical decision (Onifer & Swinney, 1981; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982). Experiment 2 shows that the relation between the two words that is reflected in lexical-decision responses is also available to give inhibition in recognition responses.

In contrast, the computed meaning of a text does not appear to be available immediately. Response times for test words were not affected by the relevance of the test word to the meaning of the text as a whole. The suggestion is that the processes of constructing a text representation must be relatively time consuming. Although information based on direct semantic association is available quickly, integration of this information with the text-representation and computation of other information for the text representation must take longer than the time available in Experiment 2. How much longer is still an open question. Experiment 2 shows that the information is not available with only one word presented for only 250 ms. Previous research (McKoon & Ratcliff, 1986, 1987) shows that the information is available after several words (250 ms each). These endpoints would seem to bracket the time course for computing the information. However, further research would be needed to determine whether it is time alone that determines when the text representation is computed or time plus some other factor such as the content of intervening words.

Although Experiment 2 appears to support the conclusion that only the semantic-associative information was available immediately, there is one possible problem, and that is that the predicted test words for the weak-association materials may not have been predictable enough to be inferred. In other words, even if there had been plenty of time to compute the relation between the predicting sentence and the test word, it may never have been computed. We did not think this was the case, because the materials used in Experiment 2 were a subset of materials previously shown to lead to inferences (McKoon & Ratcliff, 1986). But the materials were not exactly the same as those used previously, so we thought it best to demonstrate empirically in Experiment 3 that they would lead to inferences.

Experiment 3

This experiment was designed to show that the materials from Experiment 2, the predicting sentences without semantic associations, would lead to inferences if sufficient time for construction of inferences was available. The procedure used to show the inferences was the delayed item-recognition procedure from Experiment 1. On each trial, two sentences were presented for study, and then six items were presented for test. Sentences were studied in either their predicting or their control form, and the test item of interest was the word predicted by the predicting sentence. This word was always primed by a word from the studied sentence (e.g., for the pair of sentences in Table 5, *hurt* was primed by *diver*). Correct negative responses to the test words should be inhibited in the predicting condition relative to the control condition if the inference represented by the test word was inferred during reading. This difference should obtain even if the inference was encoded only minimally.

Method

Subjects. Students from an introductory psychology class participated in the experiment for course credit; there were 24 subjects total.

Materials. The materials were the 16 sentences from Experiment 2 without strong associations between predicting sentence and test word (e.g., the first pair in Table 5). There were also 32 sentences used as fillers. The procedure was exactly the same as that used in Experiment 1. The study and test lists were constructed in the same way as in Experiment 1 except that the prime for the predicted test word was always a word from the sentence and never the neutral word *ready*.

Procedure. The experimental sentences were presented in either their predicting or their control version. This variable was combined with two sets of sentences and two groups of subjects in a Latin-square design.

Results

As expected, responses for the predicted test words were less accurate when the predicting sentence had been studied than when the control sentence had been studied (35% errors vs. 21% errors, respectively). The data were analyzed in the same way as the data in Experiment 1, and the difference in error rates was significant, $F(1, 23) = 9.8$, with subjects as the random variable, and $F(1, 15) = 6.2$, with sentences as the random variable. The two conditions did not differ significantly in average response times (321 ms and 303 ms, respectively).

For filler test items, the data were as follows: positive test words primed by a word from a studied sentence, 243 ms (17% errors); positive test words primed by *ready*, 251 ms (21% errors); negative test words primed by a word from a studied sentence, 288 ms (16% errors); negative test words primed by *ready*, 275 ms (16% errors).

The results of Experiment 3 show that the inferences represented by the predicted test words were generated from the predicting sentences. Thus, in Experiment 2, when there was no effect of the inferences on responses, it could not have been because the inferences would never have been generated. Rather, the lack of effect in Experiment 2 can be interpreted as due to the lack of time available for construction of inferences.

Conclusions

In the introduction of this article, we proposed a framework that stresses variability in the degree of encoding of inferences as well as the importance of investigating the sources of this variability. The results from Experiments 1 and 2 examine the effects of semantic association on inference processes. These effects provide sufficient information to formulate theoretical statements about some of the factors involved in inference processing and, in doing so, demonstrate that progress in understanding inference can come about through parametric manipulations of variables.

According to the view that we propose for inference processing, an inference represents encoded features of the meaning of a text, features that were not explicitly stated. Three assumptions are made about inferences, each discussed in the following paragraphs: First, the sets of features that represent inferences can vary in their degree of specificity or focus; second, inferences can vary in the strength with which they are encoded; and third, features can vary in the time course of their availability. When we used the term *feature*, we mean aspects of meaning and do not necessarily mean feature as in a feature or exemplar model. In a propositional framework such as Anderson's (1983, chap. 5), aspects of meaning might be represented as propositions about the concepts; for example, if the concept *dog* was encoded, propositions that represent *big, cuddly, savage when hungry*, and so on might be encoded, depending on relevance to the text, along with propositions from the text.

The degree of specificity or focus of an inference refers to the number of different features that might represent an inference from an explicitly described event. Most of the features for an inference from the words *seamstress, needle*, and *thread* would have to do with *sewing*, and so the inference would be highly focused. The set of features that might be generated from *actress, fall from*, and *14th story* would be much more diffuse and would have to be compatible with many, more diverse, possibilities, perhaps including dying, various ways of getting badly hurt, mechanisms for miraculous rescues, and ways of summoning aid. Whatever features were encoded as an inference from *actress, fall from*, and *14th story*, they would have to be potentially compatible with all of these possible outcomes. In addition, the degree of specificity or focus of the features of an inference is related to its goodness of encoding. Because the words *seamstress, needle*, and *thread* specify the concept *sew* so exactly and completely, it may be that it could be described as explicitly instantiated. With the *14th story* example, the potential events represented by the encoded features may be so diffuse that no explicit inference is instantiated at all.

The scheme whereby explicit encoding of inferences depends on their degree of focus incorporates sensible design principles. In the *sew* example, it is unlikely that the text is about anything other than sewing. If the text followed with some other alternative, it would have to be highlighted or signaled, and possibly extra processing would be required to achieve comprehension. If the text was not about sewing, then the inference would have to be canceled in some relatively explicit way. For example, in the control sentence in Table 1, to make an inference about sewing less likely, the seamstress is careless, the needle is unthreaded, and the needle is on the floor. In contrast, in the *falling from 14th story* example; there are many possible continuations, as already mentioned. If a reader made a specific inference, it would be wrong for a large proportion of possible continuations, and backtracking and disruptions to comprehension would result. Instead, sensible design principles suggest readers make inferences that are not focused but, rather, compatible with a range of possibilities.

A second assumption about inferences is that they vary in the strength with which they are encoded. For highly focused inferences, variations in strength are easy to imagine. For example, *sew* would be a highly focused and strong inference for the sentence in Table 1 about the seamstress threading a needle. But if the seamstress did not sew (because in the next sentence, the lights went out), a set of features might still be encoded that focused on *sew*, but they might be encoded much more weakly. For diffuse inferences, inferences which have a large range of compatible next events, degree of focus is not so easily separated from strength because diffuse inferences may necessarily be weak inferences. However, it may be that there are variations in strength that go from weak to very weak, or from weak to no inference at all.

The third assumption of our view of inference processing is that the features of meaning that make up inferences vary in the time course of their availability. Some are available quickly, directly from memory, whereas others require computation. The features of meaning that connect *seamstress*, *needle*, and so on to *sew* would be immediately available because there are direct associations among them in memory. Associations that are not directly available from memory will require time for computation. A highly focused inference can become available relatively quickly or more slowly, depending on whether the features of meaning required for the inference are directly available from memory or must be computed. A more diffuse inference would require time for computation if the features making up the diffuseness were not directly available from memory. However, it might be argued that a range of features could all be directly available from memory. For example, a word presented out of context might be encoded with a range of variations on its meaning, all directly and immediately available.

These three assumptions about inference processing combine to provide an account of the results of the experiments in this article as well as the results of much previous work. With respect to the current work, the assumptions together can be used to interpret the data from Experiments 1 and 2. In Experiment 2, with the strong-association materials, the features of meaning that connect the final word of a predicting sentence to the test word are immediately available because there is a direct association in memory between the final word of the sentence and the test word. The features of meaning are also relatively well focused because *sitting* is by far the most typical thing to do with *chairs*, given the context of the sentence. With the weak-association materials, the features that would be encoded from the predicting sentence relative to the test word would be more diffuse and would require time for computation. *Hurt* is not immediately available from *cement*, and so responses to *hurt* are not inhibited.

Another case of features requiring time for computation is illustrated by the nonpredicting control sentences with strong associations. The inference that might be made from the sentence about shopping and finally finding the perfect chair would be something about buying the chair, not sitting in it. Computation of features that specified this inference and computations to cancel the features having to do with sitting might take some time, whereas features representing the direct association between *chair* and the test word *sit* would be available more quickly. Thus, the data show inhibition for *sit* even when the sentence predicts buying a chair, not sitting in a chair.

Experiment 1 was designed to measure the goodness of match between a text in memory and a test word representing a possible inference from that text (cf. Ratcliff & McKoon, 1988a, 1988b). The goodness of match should have a relatively high value if the features of the inference were highly focused on the concept represented by the test word and the inference is strongly encoded. The goodness of match should have a relatively low value if the features were more diffuse and not centered on the concept represented by the test word. This pattern is what the data show; for test words with direct semantic associates in the text, the match between test word and text is good enough to give significant inhibition on responses to the test word, about as much inhibition when the test word is presented by itself with only the neutral prime as when the test word is presented with a prime from the text. In contrast, the match between test words with weak associations to their texts was not as good. When these test words were presented with the neutral prime, there was not as much inhibition as with a prime from the text.

For the strong-association materials in Experiment 2, note that inhibition is not determined by the match between the individual words of the text in memory and the test word. These words are the same in the predicting and control conditions. Instead, inhibition is determined by the relation between the test word and the text as a whole, so that inhibition is greater in the predicting than in the control conditions. In other words, the effect of the semantic associates in the text is not to determine inhibition directly but, rather, to determine the inference that is made from the text, and it is the inference that determines inhibition.

Besides the data from the experiments reported in this article, the assumptions about the degree of focus and strength of inferences and about the time course of inference processing fit well with many previous results. For example, direct semantic-associative information has been shown to be immediately and automatically available in word-to-word priming in lexical decision (Meyer & Schvaneveldt, 1976) and in sentence-to-word priming (Kintsch & Mross, 1985; Onifer & Swinney, 1981; Seidenberg et al., 1982). At the same time, in these paradigms, information that does not derive from direct semantic associations takes longer to compute (Glenberg et al., 1987; Neely, 1977; Onifer & Swinney, 1981; Seidenberg et al., 1982).

Another area of research in which semantic associations might be assumed to play a role is the area concerned with inferences about the meanings of words with respect to their contexts (Barsalou, 1982; Tabossi, 1982). Typically, subjects are asked to verify features of target words presented in sentence contexts. For example, following a sentence about painting a picture of a tomato, they might be asked to verify the sentence "Tomatoes are red." When the verification sentences are presented immediately after the context sentences, so that both would be in working memory at the same time, then features more closely associated to the target word lead to faster response times (Barsaolu, 1982; Tabossi, 1982). However, this result must be qualified by aspects of the experimental procedure that might have led subjects to guessing strategies (see McKoon & Ratcliff, 1986). In other experiments (McKoon & Ratcliff, 1988b) where strategic processing could be ruled out, there is evidence that contextually appropriate features of words are inferred and encoded into memory with the context. But such evidence is found only in some retrieval environments, suggesting that the features are encoded weakly and suggesting an analogy with the weak-association materials in the experiments presented in this article (see also Ratcliff & McKoon, 1988a). How the strength of semantic associations between potentially encoded features of meaning of words and a text might affect the strength of encoding is a question that has not yet been addressed empirically.

Another finding compatible with the view proposed in this article is the large effect that encoding manipulations can have on elaborative inferences. For example, evidence that readers infer the instruments of verbs appears only when subjects are explicitly told to make such inferences (Dosher & Corbett, 1982) or when the instrument is highly semantically associated to the verb and it is mentioned prior to the verb in the text (McKoon & Ratcliff, 1981). To give another example, readers connect two different stories that have the same theme only when they are given specific instructions and tasks that require them to do so or when they are given sufficient exposure to one of the stories to make it very well learned (Seifert, McKoon, Abelson, & Ratcliff, 1986). These findings are consistent with the notion that elaborative inferences based on weak associations may only be computed under special circumstances, or without special circumstances they may only be computed minimally.

The emphasis in this article on semantic association as a determinant of degree of specificity of inferences fits well with a more general view of text comprehension that we have suggested previously. This view represents a summary of research in text processing according to which inferences establish mainly local coherence with only minimal encoding of other kinds of inferences (McKoon & Ratcliff, 1988b). In McKoon and Ratcliff (1986), we distinguished inferences necessary to connect propositions by argument repetition from inferences not necessary to achieve coherence. In the first category are the inferences that establish connections between two instances of the same concept (McKoon & Ratcliff, 1980b; Ratcliff & McKoon, 1978) and the inferences that establish the referent of an anaphor (Haviland & Clark, 1974; Corbett, 1984; Corbett & Chang, 1983; Dell, McKoon, & Ratcliff, 1983; McKoon & Ratcliff, 1980a). Evidence that these inferences are generated quickly and automatically is abundant, and we would interpret this as an indication that information in the working memory representation of a text (Kintsch & Vipond, 1979) is, like direct semantic-associative information, easily available. In the second category are inferences that are not necessary for coherence, such as the elaborative inferences studied by McKoon and Ratcliff (1986) and Singer and Ferreira (1983), inferences that fill in schema information (Alba & Hasher, 1983; Seifert et al., 1986), and inferences about the instruments of verbs (Corbett & Dosher, 1978; McKoon & Ratcliff, 1981). Also in this category may be inferences about the global structure of a text. For all of these elaborative inferences, there is little evidence that they are automatically processed during reading. This follows from the view presented in this article, that without the support of directly available information, explicit encodings of these inferences require more computation than would usually be performed.

References

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

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

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

Corbett, A. T. (1984). Pronominal adjectives and the disambiguation of anaphoric nouns. *Journal of Verbal Learning and Verbal Behavior, 17*, 479–491.

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.

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.

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

Forster, K. (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.

Garrod, S., & Sanford, A. J. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior, 16*, 77–90.

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

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.

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.

Kintsch, W., & Mross, E. F. (1985). Context effects in word identification. *Journal of Memory and Language, 24,* 336–349.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. Nilsson (Ed.), *Perspectives on memory research* (pp. 329–365). Hillsdale, NJ: Erlbaum.

McKoon, G. (1988). *Word identification and elaborative inference: The relation between theory and empirical measurement.* Unpublished manuscript.

McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 463–480.

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. (1987, November). *Semantic association and inference processes.* Paper presented at the meeting of the Psychonomic Society, Seattle, WA.

McKoon, G., & Ratcliff, R. (1988a). *Assessing the occurrence of elaborative inference with recognition: A reply to Potts, Keenan, and Golding.* Unpublished manuscript.

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

McKoon, G., & Ratcliff, R. (1988c). *Inferences about contextually-defined categories.* Unpublished manuscript.

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science, 192,* 27–33.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity intention. *Journal of Experimental Psychology: General, 106,* 226–254.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9,* 225–236.

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. (1981). Automatic and strategic priming in recognition. *Journal of Verbal Learning and Verbal Behavior, 20,* 204–215.

Ratcliff, R., & McKoon, G. (1988a). Memory models, text processing, and cue-dependent retrieval. In F.I.M. Craik & H.L. Roediger (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving.* Hillsdale, NJ: Erlbaum.

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

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

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

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.

Singer, M. (1978, August). *The role of explicit and implicit recall cues.* Paper presented at the meeting of the American Psychological Association, Toronto, Ontario, Canada.

Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 22,* 437–448.

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

Received July 13, 1987
Revision received March 21, 1988
Accepted May 16, 1988 ■