

Availability and Accessibility of Information and Causal Inferences From Scientific Text

Jennifer Wiley

*Department of Psychology
University of Illinois at Chicago*

Jerome L. Myers

*Department of Psychology
University of Massachusetts, Amherst*

In these experiments, we examine inferences from simple scientific passages that provided all inference-relevant information for the reader. In Experiment 1, we use a contradiction paradigm to test whether readers compute causal bridging inferences from scientific text. In Experiments 2 and 3, we use the same methodology but vary the availability of the critical information by manipulating the distance between critical pieces of information. The results of the experiments indicate that causal bridging inferences can be computed from scientific text when all the necessary information is available and accessible.

Understanding the meaning of a text usually requires the reader to draw inferences, that is, to determine the message that is intended by the writer, even though the meaning is not made fully explicit in the text itself. A major issue concerning many text researchers has been the kinds of inferences people usually make in the course of comprehending a text, and when, why, and how these inferences are made. There has been considerable debate about the answers to these questions, and several inference types have been proposed in order to try to delineate which inferences may be more basic, necessary, obligatory, or even automatic in the course of text processing, as opposed to which may be more

elaborative, voluntary, or strategic. Distinctions have been made between text-based and knowledge-based inferences, bridging and elaborative inferences, logical or deductive inferences versus probabilistic or inductive inferences, and inferences that are made routinely as opposed to inferences that are made only under specific conditions or in narrow time windows (e.g., Graesser, Singer, & Trabasso, 1994; Just & Carpenter, 1980; Kintsch, 1988; Lea, 1995; Potts, Keenan, & Golding, 1988). In each case, the former category is intended to capture “necessary” inferences, whereas the latter refers to inferences that are possible, but not needed for initial comprehension.

The general consensus from prior studies is that the kinds of inferences that are made online are those that are necessary to maintain the coherence of the text representation. Most studies indicate that inferences that establish coherence by building connections or relations between a current text proposition and a prior text proposition are made online (e.g., Graesser et al., 1994; Haviland & Clark, 1974; Just & Carpenter, 1980; Lea, 1995; McKoon & Ratcliff, 1980; Paris & Lindauer, 1976; Potts et al., 1988; Singer, 1988), whereas inferences that go beyond relating parts of text are seen as elaborations, and are made only under specific constraining conditions, such as when readers have the required knowledge or particular reading goals (Murray, Klin, & Myers, 1993; O’Brien, Shank, Myers, & Rayner, 1988), or may be only partially activated (Duffy, 1986; McKoon & Ratcliff, 1986, 1992) or activated for limited amounts of time but not durably encoded (Cook, Limber, & O’Brien, 2001). There is general agreement that readers spontaneously generate causal bridging inferences when reading narrative text (e.g., Klin, 1995; Potts et al., 1988; Suh & Trabasso, 1993; van den Broek, 1990); however, there is considerably less evidence that similar inferences are necessarily drawn from expository texts.

Evidence from narrative text experiments using both reading time and probe naming methodologies suggests that causal information from earlier in a narrative text is reinstated at causal coherence breaks, and that readers quickly make causal inferences in the absence of an explicitly stated cause in the text (i.e., Klin, 1995; Suh & Trabasso, 1993). The reinstatement of causal information is predicted by both constructionist models (Graesser et al., 1994), which assume that readers seek a cause for each event or action in a text, and by memory-based text processing models (McKoon & Ratcliff, 1992; Myers & O’Brien, 1998), which assume that information related to the current input is passively reactivated. In view of both agreement between opposing theoretical positions and the evidence from narrative text studies, it is somewhat surprising that evidence for the generation of causal bridging inferences is less clear when comprehension of expository text is considered. The goal of the present investigation is to attempt to reconcile some contradictory findings in relation to causal bridging inferences from expository text, to determine if casual bridging inferences are computed as part of scientific text comprehension, and if so, to determine when and how they are computed.

EVIDENCE FOR CAUSAL BRIDGING INFERENCES IN EXPOSITORY TEXT

There has been much less research using expository or scientific texts than there has been using narrative texts, and no clear consensus on the status of causal inferences has emerged from these studies. Noordman, Vonk, and Kempf (1992) failed to find evidence for online scientific causal inferences, except when inferential strategies were explicitly encouraged. On the other hand, Millis and Graesser (1994) and Singer and his colleagues (Singer & Gagnon, 1999; Singer, Harkness, & Stewart, 1997) have concluded that causal inferences are generated by readers of scientific text without explicit encouragement to do so.

There are several key observations about the methodologies used in the studies of Noordman et al. (1992), Millis and Graesser (1994), and Singer et al. (1997) that need to be noted before their results can be compared, and the present methodology described. First, there are two main factors on which these studies differ—the difficulty of the content and the method used to test for inferences. Noordman et al.'s passages were quite difficult and on unfamiliar science topics, whereas Millis and Graesser's study and Singer's studies used topics that were more accessible. Secondly, Millis and Graesser used a lexical decision time methodology, in which inference generation was measured in terms of facilitation of responses to inference-related probe words. In the Noordman et al. and Singer et al. experiments, reading times on target sentences containing an inference were measured.

The passages in the Noordman et al. (1992) study contained a target sentence having a conclusion/premise pair, connected by the word "because" such as "Silicon was used for the space shuttle because weight is a factor." In half of their passages, the target sentence was preceded by a sentence containing another premise such as "Silicon is light." That premise was omitted on the remaining half of the trials. Noordman et al. reasoned that if readers generate causal bridging inferences online, then they will compute the missing premise as soon as they are presented with the target sentence (the conclusion/premise pair). If readers do infer the missing premise (i.e., the implication of the conclusion/premise pair), this process should require time. Therefore, reading times should be longer on the target sentence when it is not preceded by the premise "Silicon is light." On the other hand, if the implicit premise is not computed when the reader encounters the target sentence, reading times on the target sentence will not differ between conditions. Furthermore, Noordman et al. reasoned that if implicit premises are computed by readers, verification times on those implicit premises presented after the target sentence should not differ between conditions. However, if the implicit premise is not computed during reading, verification times on that premise will be longer in the implicit condition, as the reader will need to verify the missing premise when it is read. Noordman et al. found that when a premise was left implicit,

readers did not take significantly longer to read a consistent conclusion, but they did take longer to verify the missing premise. This pattern of results, using both conclusion reading and premise verification times, was taken as an indication that readers did not infer the missing premise spontaneously.

A subsequent experiment demonstrated that readers in the implicit condition did take longer to read the "because" sentences when they were told they would have to answer questions such as "Why did NASA use silicon?" Readers in the implicit condition also had longer reading times on the "because" sentence when they were told to judge whether each sentence was consistent with previous portions of the text. These data suggest that although readers did not attempt to make bridging inferences spontaneously while reading these expository texts, such inferences could be made when the reader's goal required it, as indicated by the longer reading times on the conclusion sentences. (Verification probe times are not reported for this condition.)

Although these data are suggestive, one interpretation of these results is that the reading measures provided here do not yield a sensitive test of inference generation. Reading time on the "because" sentence could be fast, either because readers have all the necessary information to compute a premise and can make an inference quickly and effortlessly, or because they are reading superficially, and failing to compute the premise, construct the inference, or verify a relationship to a previous sentence, especially if they lack prior knowledge about the topic. More problematic is the interpretation of the verification times. Verification of the "missing" premises could have been facilitated in the explicit condition simply because they were actually presented in the text. Therefore, there are alternative explanations for the shorter reading and verification times in the explicit condition and they may not reflect whether or not an inference has been computed.

In contrast to the Noordman et al. (1992) results, Millis and Graesser (1994) obtained evidence that causal inferences are made spontaneously from expository text. In their study, inferences were investigated using a lexical decision task. Following each sentence in an expository passage, a letter string was presented, and participants had to respond "yes" or "no" to indicate whether or not the string was a word. An example passage about nuclear power is included in Table 1. The sentences are on the left and the probe string, in capital letters, is on the right. Readers made faster lexical decisions when the probe word was related to a causal antecedent inference (RISES) from that passage than when the probe word was presented at the end of a sentence in another unrelated passage. On the basis of these findings, Millis and Graesser concluded that readers do spontaneously generate causal antecedent inferences online from expository, scientific texts.

One obvious difference between the Noordman et al. (1992) and Millis and Graesser (1994) experiments is that the passages in the Noordman et al. study were more complex, and on highly technical topics, whereas Millis and Graesser used five-sentence passages on more familiar topics, with few words and clear

TABLE 1
Example Passages From Studies of Expository Text

Millis & Graesser (1994; RSVP procedure, lexical decision probes in capital letters on right)		
NUCLEAR POWER		
Neutrons split atoms into two particles.	KEND	
Heat energy is released.	WATER	
The water in a surrounding tank is heated.	VIERCE	
Steam drives a series of turbines.	RISES	(inference)
The turbines produce electricity.	STALPH	
Are neutrons heated to release energy?		
Singer & Gagnon (1999; reading times were measured)		
A lubricant is used to prevent contact of parts in motion and thereby reduce friction.	<i>Neutral</i>	
Lubricants should be good coolants.	Premise	
Liquid hydrocarbons are the most commonly used lubricants	Conclusion	
because they effectively remove/add heat.	Premise	
A lubricant may also prevent rusting of close fitting parts.	Neutral	
Are lubricants good coolants?	Question	

points. Noordman et al.'s passages may have taxed both the knowledge and the comprehension ability of students. This may have limited the ability of novice readers to create coherent representations of the texts. Results obtained by Singer et al. (1997) indicated that simplifying and clarifying the Noordman et al. passages may indeed facilitate the construction of causal inferences.

A second major difference is in the task that was used to assess inference generation. Although, as previously noted, there are potential problems in interpreting reading time data on consistent conclusions, lexical decision tasks have their problems as well. As with any probe task, lexical decision tasks carry the risk that any observed effects may reflect semantic associations between the text and the probe word, rather than facilitation due to the computation of an inference. In addition, tasks that require a binary response, such as the lexical decision task, may yield facilitation effects not as the direct result of the generation of an inference, but simply as an artifact of the compatibility of the probe with the preceding text, which may speed the decision (Potts et al., 1988; Seidenberg, Waters, Sanders, & Langer, 1984; Wiley, Mason, & Myers, 2001). Especially when decision-based probe results have long response times, as was the case with the Millis and Graesser (1994) study, there is a chance that any observed effects are not due to differences in priming of concepts, but differences in a postaccess decision phase, during which context checking may play a role.

On the other hand, Singer and Gagnon (1999) have found evidence that is consistent with the findings and conclusion of Millis and Graesser (1994). Using a contradiction paradigm (see Table 1), they adapted passages from Noordman et al.

(1992) and created an additional set of simpler scientific passages. Readers either saw the accurate, consistent version of the target sentence, "Liquid hydrocarbons are the most commonly used lubricants because they effectively remove heat," or they saw an inconsistent version, "because they add heat." The substitution of inconsistent conclusions into passages, and evidence of their detection in reading times, is the basis of the "contradiction paradigm." This method has been an effective approach for identifying inference-related processing in narrative text (e.g., Albrecht & Myers, 1995; Albrecht & O'Brien, 1993; Klin, 1995; Myers, O'Brien, Albrecht, & Mason, 1994). The logic is that if the reader has drawn an inference, then a line containing a conclusion inconsistent with that inference will be read more slowly as the reader attempts to resolve the inconsistency. Indeed, in the Singer and Gagnon study, readers took longer to read the inconsistent version of the target "because" sentence, indicating that they noticed the contradiction. These results suggest that readers can compute causal inferences from scientific text.

The Singer and Gagnon (1999) experiments address several potential issues with the Noordman et al. (1992) study by using a contradiction paradigm and simpler passages. However, there were a few aspects of the Singer materials that may still be prompting readers to compute inferences that they might not make otherwise. Because both the conclusion and the second premise were presented in the same target sentence, and reading times were collected for the complete sentence, we do not know if the generation of the inference depended on both being presented simultaneously. Readers may have drawn the inference ("they effectively remove heat") immediately after reading the conclusion, or may have been prompted to validate the conclusion after they were presented with the cause. In either case, the generation or verification of the inference could have depended on the premise and conclusion being presented in close temporal succession. Further, we do not know what role explicitly marking the causal inference with the connective "because" played in the noticing of the contradiction. Finally, the comprehension questions used in the Singer study directly tapped the causal inference that could be made in each passage, which may have prompted readers to adopt a strategic reading style. The comprehension questions could have had an effect similar to Noordman et al.'s reading goal manipulation. Thus, although the results of Singer and Gagnon's experiments indicate that inferences can be drawn by readers of expository text under one condition, they do not tell us whether readers will compute causal inferences spontaneously or when such inferences may be generated.

In these experiments, we separated the premises from the causal inference conclusions in short scientific texts. Instead of presenting a premise, and then a conclusion and second premise in a sentence connected with a "because," we presented readers with one or two premises first in separate, unmarked sentences followed by an unmarked conclusion. We measured the time to read a line presenting a conclusion that was either consistent or inconsistent with the premises stated in preceding sentences.

In Experiment 1, the premises immediately preceded the conclusion and in Experiments 2 and 3, the availability of the critical information was varied by manipulating the distance between premises, and between the two premises and the conclusion. When two premises precede the conclusion, readers have the basis for forming an inference. Therefore, if readers spontaneously draw causal bridging inferences while reading expository text, they should read an inconsistent conclusion more slowly than a consistent one. However, when only a single premise is present in the text, there are not sufficient grounds for an inference, and no contradiction effect is expected.

EXPERIMENT 1

In several studies of narrative text processing, it has been shown that participants read a sentence more slowly when it contains a conclusion that was inconsistent with an implicit proposition. For example, Albrecht and O'Brien (1993) found that readers "boggled" when reading a proposition that was inconsistent (Linda ate a cheeseburger) with an earlier proposition (Linda was a vegetarian). Reading time on the target sentence in inconsistent conditions was compared to the same sentence in conditions in which readers were never told that Linda was a vegetarian. Using scientific texts, however, Noordman et al. (1992) found that readers did not routinely notice contradictions. This is consistent with a number of findings that readers of expository text are generally insensitive to inconsistencies (Glenberg, Wilkinson, & Epstein, 1982; Markman, 1979; Otero & Kintsch, 1992). In contrast, Singer and Gagnon (1999) found evidence that readers of expository text can notice causal inconsistencies in scientific text.

Singer and Gagnon (1999) presented evidence that readers can compute causal inferences as they read expository text. However, as previously noted, their design leaves open the questions of what role the causal markers or inference-related comprehension questions may have played, as well as the issue of whether presenting the conclusion and second premise together was in part responsible for the differences in reading times. To address these questions, we performed a similar study measuring reading time on target lines containing conclusions. In our passages, both the premises and the conclusion were presented on separate lines, with no causal signals like "because." Furthermore, comprehension questions were not related to the target inferences. A complete example passage is included in Table 2.

We classify the inferences that we are examining as causal bridging inferences. Because the conclusion that follows the two premises relates to the cause of a particular scientific phenomenon, we have categorized these conclusions as causal inferences. Assuming a conclusion is the result of integrating the two premises stated in the text, we follow the usual practice of categorizing the inference as a bridging inference. In addition, since all necessary information for the inferential

TABLE 2
Example Passages From Experiment 1

Two Premise/Consistent Version	
SEALS	
Seals are usually found in cold regions.	
Like most animals in such regions, they usually have to produce	Premise 1
a lot of energy just to keep warm.	Premise 1
<i>Metabolic rate increases with energy needs.</i>	<i>Premise 2</i>
Seals have high metabolic rates.	Conclusion
They generate energy from a diet of mollusks and fish.	
Are seals usually found in tropical regions?	Question
One Premise/Inconsistent Version	
SEALS	
Seals are usually found in cold regions.	
Like most animals in such regions, they usually have to produce	Premise 1
a lot of energy just to keep warm.	Premise 1
<i>Metabolic rate is affected by energy needs.</i>	<i>Non-premise Control</i>
Seals have low metabolic rates.	Conclusion
They generate energy from a diet of mollusks and fish.	
Are seals usually found in tropical regions?	Question

conclusion is spelled out for the reader, these inferences can also be considered deductive in form. This represents a difference between this study and both the Singer and Noordman studies, which presented readers with only one premise before the conclusion. However, by presenting readers with all necessary information, we are in a position to test whether causal bridging inferences may be computed from scientific text in a condition where prior knowledge is much less likely to be a factor.

If readers are able to spontaneously compute the inference after they are presented with both of the premises, then reading times should be longer on the inconsistent than on the consistent conclusion (the bold line in Table 2). In the absence of one of the premises (the italicized line), readers of these passages (who we presume have little prior knowledge about the specific topics) should not be able to compute the inferences, and thus should show no effects of “consistent” versus “inconsistent” conclusions. Thus, an interaction between consistency and number of premises in the reading times on the conclusion lines would be evidence that causal bridging inferences are generated when both premises are available.

Method

Participants. Thirty-two undergraduate psychology students attending the University of Massachusetts at Amherst participated in this experiment for extra credit, as part of the Psychology Department subject pool.

Materials. The 24 experimental and 24 filler passages for this study were adapted from the Childcraft “How and Why” series of children’s encyclopedias (1989), as well as from prior studies of scientific expository texts (Cote & Goldman, 1999; Glenberg et al., 1982; McNamara, Kintsch, Songer & Kintsch, 1996; Meyer, 1985; Millis & Graesser, 1994; Noordman et al., 1992). Two versions of an example passage are presented in Table 2. All passages were between 4 and 6 sentences, and between 6 and 11 lines in length. Each line typically contained between 5 and 10 words, and was no longer than 50 characters. Each passage contained an embedded syllogism, such as: *Seals have to produce a lot of energy. Metabolic rate increases with energy needs. Seals have high metabolic rates;* or *Himalayan ice is less brittle than Alpine ice. When a glacier breaks, it creates an avalanche. Alpine avalanches are more frequent than Himalayan avalanches.* Each passage started with a title line in capital letters. The first few lines introduced the topic. Following the introductory lines, the next few lines contained the 2 premises. The line after the premises contained either a consistent conclusion (Seals have high metabolic rates) or an inconsistent conclusion (Seals have low metabolic rates). Following the conclusion line, a final sentence ended the passage. The last sentence continued the overall topic but did not relate directly to the inference sentence. It was intended to serve as a neutral post-target sentence. The consistency of the target line was manipulated by changing a single word in the conclusion (generally an opposite word such as good/poor or higher/lower). The order of the major and minor premises was not controlled. For 16 passages, the minor premise (containing the subject of the conclusion) preceded the major premise (containing the predicate); for the other 8, the major premise came first. Post hoc analyses showed identical patterns for both orders.

Four versions of each passage were created. Two had consistent conclusions and two had inconsistent conclusions. The conclusion line is indicated in bold in the examples. Within each consistency condition, one version contained both premises and a consistent conclusion and the other contained only one premise. The sentence containing the other premise was replaced by a neutral filler sentence (the “nonpremise control”) of approximately the same length and number of syllables. In the examples in Table 2, the premise that is replaced and the nonpremise control are indicated in italics. The assignment of items to conditions was counterbalanced across four lists.

In addition, there were 24 nonexperimental passages randomly interleaved among the experimental passages. These filler passages were on scientific topics, and were of the same length as the experimental passages, but they were largely descriptive and did not require the reader to draw any causal inferences. They also did not contain any contradictions.

A yes/no comprehension question followed each passage. The question referred to information that was explicitly stated in all versions of each passage, and did not relate to the conclusion line. Half of the experimental and nonexperimental passages required a “yes” response, and half of each required a “no” response.

Procedure. All materials were presented on a monitor controlled by a 386 computer. Participants read the passages in a self-paced manner, advancing the text one line at a time by using a response lever. Participants were instructed to read carefully so that they would comprehend the passages and be able to answer questions about the passages after they had read them. Each passage was followed by a yes/no question that did not relate to the inference.

Results

In all analyses, F_1 refers to tests against subject variability and F_2 refers to tests against item variability. In Experiments 1 through 3, times of more than two standard deviations from each participant's mean reading time on the conclusion line were not included in the analyses. Less than 5% of the scores was omitted by this criterion. All p values are less than .05 unless otherwise reported.

Mean reading times for the line containing the conclusion are presented in Table 3. The mean reading time was significantly longer for inconsistent than for consistent conclusions; $F_1(1, 28) = 13.6$, $MSE = 133,515$; $F_2(1, 23) = 8.62$, $MSE = 166,406$. Reading times on conclusions when both premises were present were longer than when only one was present. This effect was significant only when tested against subject variability; $F_1(1, 28) = 5.0$, $MSE = 126,850$, $p < .03$; $F_2(1, 23) = 1.6$, $MSE = 139,848$, $p = .21$. Both main effects seem to be a result of the markedly longer time in the inconsistent/two-premise cell. This is reflected in the interaction between number of premises and consistency of conclusion, which was significant by both subjects and items, $F_1(1, 28) = 7.48$, $MSE = 170,607$; $F_2(1, 23) = 5.41$, $MSE = 110,385$. Planned comparisons indicated that when both premises were present, the line containing the inconsistent conclusion was read much more slowly than the line containing the consistent conclusion ($F_1(1, 28) = 15.47$; $F_2(1, 23) = 12.73$), whereas there was little difference between the means when one premise was absent; both $F_s < 1$.

Readers detected an inconsistency only when they had all the information needed to compute the inference. They were not able to recognize inconsistent conclusions in the single premise condition, when they were required to rely on their own knowledge to recognize the inconsistent conclusion. However, the fact that the inconsistency of the conclusion was detected when both premises were

TABLE 3
Mean Reading Times (in ms) on Conclusions by Consistency and Number
of Premises for Experiment 1

	<i>Consistent</i>	<i>Inconsistent</i>	<i>Difference</i>
Both premises	1967	2405	438
One premise	2025	2064	39

presented indicates that when readers are given all necessary information explicitly in scientific passages, they are able to compute inferences. This extends Singer & Gagnon's (1999) results by showing that scientific causal inferences are computed even when the conclusion and final premise are presented without a causal connective such as "because," and are computed without the encouragement afforded by questions targeting the inference.

One possibility is that causal inferences are computed only to the extent that the information needed to draw the inference is easily available to the reader. In Experiment 1, both premises immediately preceded the conclusion and are likely to have been in working memory. Therefore, Experiment 2 was designed to delineate the conditions under which these inferences may be drawn by manipulating the accessibility of the critical information in the text.

EXPERIMENT 2

Although in previous experiments readers have often failed to notice contradictions in expository text (Glenberg et al., 1982; Noordman et al., 1992; Otero & Kintsch, 1992), we found significant increases in reading time on inconsistent conclusions when students had all the required information to make inferences from short scientific texts, consistent with the findings of Singer and Gagnon (1999). Why were readers in these experiments able to recognize contradictory statements so well, while other studies have failed to find contradiction effects? In the case of our study, the passages had two key elements. First, readers had all the necessary information provided to them explicitly in the text. Secondly, similar to the Singer and Gagnon study, the relevant information was presented in close proximity to the conclusions. The proximity of premises to the conclusion may have been responsible for the accessibility of the causal information, and the generation of causal inferences. To test whether proximity affects the generation of causal inferences, filler material was inserted between the critical pieces of information in the passages. To accomplish this, but keeping the actual information presented constant across experiments, we moved part of the introductory lines to later locations in the passage. For each passage, the lines were inserted either between the two premises or between the premises and the conclusion. We were interested in whether causal inferences would be less likely to be made as the information required for such inferences became less accessible.

Method

Participants. Thirty-two undergraduate psychology students attending the University of Massachusetts at Amherst participated in this experiment for extra credit, as part of the Psychology Department subject pool.

Materials. The 24 experimental passages for this study were adapted from the first experiment. Examples of each distance condition are presented in Table 4. A single “distanced” version of each passage was created by placing part or all of the original first sentence either between the two premises or between the two premises and the conclusion. Any anaphoric or referential problems were corrected (i.e., pronouns always followed proper nouns) so that reordered passages were coherent. In a few cases, the order of the premises was reversed, to make for more natural sounding text once the filler had been inserted. Half the passages were distanced in each way—making the distance factor between items and within subjects in this experiment.

As in the first experiment, the consistency of the target line (the line containing a conclusion) was manipulated by changing a single word (generally opposite words such as good/poor or higher/lower).

Four versions of each “distanced” passage were created. Two had consistent conclusions and two had inconsistent conclusions. Within each consistency condition, one version contained both premises and a consistent conclusion and the other contained only one premise. The sentence containing the other premise was replaced by a nonpremise control sentence of approximately the same length and number of syllables. The assignment of items to conditions was counterbalanced across four lists. The same 24 nonexperimental passages were used as in Experiment 1.

Procedure. The procedure was the same as in Experiment 1.

TABLE 4
Example Passages From Experiment 2

Filler Between Premises—Two Premise, Consistent Version

SEALS

Metabolic rate increases with energy needs.

Premise 1

Seals are usually found in cold regions.

Filler

Like most animals in such regions,
they usually have to produce
a lot of energy just to keep warm.

Premise 2

Premise 2

Seals have high metabolic rates.

Conclusion

They generate energy from a diet of mollusks and fish.

Are seals usually found in tropical regions?

Question

Filler After Premises—One Premise, Inconsistent Version

SEALS

Metabolic rate is affected by energy needs.

Nonpremise Control

Seals have to produce

Premise (2)

a lot of energy just to keep warm,

Premise (2)

like most animals that are

usually found in cold regions.

Filler

Seals have low metabolic rates.

Conclusion

They generate energy from a diet of mollusks and fish.

Are seals usually found in tropical regions?

Question

Results

Mean reading times for the line containing the conclusion are presented in Table 5. As in Experiment 1, reading times were analyzed with a within-subjects 2×2 (number of premises \times consistency of conclusion) ANOVA for each distancing condition.

For the first distancing condition (where filler was inserted between the first and second premises, but the second premise was still adjacent to the conclusion), the main effect of consistency was significant by items, with reading time on inconsistent conclusions being longer than reading time on consistent conclusions, $F_1(1, 28) = 3.15$, $MSE = 472,159$, $p < .09$; $F_2(1, 8) = 4.96$, $MSE = 100,926$, $p < .06$. The main effect of number of premises was significant by subjects, but not for items, $F_1(1, 28) = 22.33$, $MSE = 228,581$, $F_2 < 1$.

The most important result for our investigation is the interaction between number of premises and consistency of conclusion, which was significant by subjects and items, $F_1(1, 28) = 30.74$, $MSE = 137,480$; $F_2(1, 8) = 5.51$, $MSE = 35,301$. As in the first experiment, planned comparisons indicated that when both premises were present, conclusion lines that were inconsistent were read much more slowly than consistent conclusion lines ($F_1(1, 28) = 15.99$; $F_2(1, 23) = 14.50$), whereas there was little difference between the means when one premise was absent; both $F_s < 1.3$. This is the same pattern as was observed in Experiment 1. These results are consistent with results found by Singer and Gagnon (1999) who performed a similar distancing experiment.

For the second distancing condition (in which filler was inserted after the two premises, but prior to the conclusion), neither the main effect of consistency nor number of premises approached significance ($F_s < 1.3$). The interaction did approach significance, $F_1(1, 28) = 3.43$, $MSE = 214,598$, $p < .08$, $F_2(1, 8) = 3.96$,

TABLE 5
Mean Reading Times (in ms) on Conclusions by Consistency and Number
of Premises by Distancing Condition in Experiment 2

	<i>Filler Between Premises</i>		
	<i>Consistent</i>	<i>Inconsistent</i>	<i>Difference</i>
Two Premises	2078	2461	383
One Premise	2130	2296	166
	<i>Filler After Premises</i>		
	<i>Consistent</i>	<i>Inconsistent</i>	<i>Difference</i>
Two Premises	2069	2053	-16
One Premise	2175	1945	-229

$MSE = 21,512$, $p < .08$, but with a very different pattern. Planned comparisons failed to reach significance when both premises were present, $F_s < 1$. However, when only one premise was present, readers tended to take more time on the consistent conclusion, although this result was not reliable ($F_1(1, 28) = 2.56$, $MSE = 360,558$, $p = .12$; $F_2(1, 8) = 2.84$, $MSE = 95,756$).

The differences between the two distancing conditions suggests that the proximity of the causal information may affect whether or not a contradiction is noticed, and presumably, whether or not the causal inference is computed, available, or instantiated. Putting filler between the two premises seemed to have little effect on the boggle; that is, the inconsistent conclusion was still noticed and took longer to read. On the other hand, inconsistent conclusions were not detected when the filler was placed after both premises and before the conclusion. Experiment 3 attempts to replicate these findings with a within-items design, in which each passage appears in each distancing condition.

EXPERIMENT 3

Experiment 3 was designed to replicate the findings of Experiment 2, and to directly test the hypothesis that the placement of filler material impacts the generation or availability of an inference, by creating multiple versions of each passage. The results of Experiment 2 are qualified by the fact that each passage was distanced in only one way. For Experiment 3, each of the original 24 passages was rewritten so that each passage had filler inserted between the two premises in one version, and inserted after the premises and prior to the conclusion in a second version. Thus, the distancing manipulations are the same as in Experiment 2, but are now fully varied across passages.

Method

Participants. Two Psychology Department subject pools were used for this experiment: 32 undergraduate psychology students attending Washington State University, Vancouver, and 32 undergraduate psychology students attending the University of Illinois at Chicago. (Half of each population was in each distancing condition; the same pattern of means was observed in each population.)

Materials. The experimental passages for this study were adapted from the first experiment. Two distanced versions were created for all passages. As in Experiment 2, one "distanced" variation of each passage was created by placing lines between the two premises. A second "distanced" variation was created by putting lines after the two premises and before the conclusion. As in previous experiments, consistency and number of premises were manipulated as within-subjects

variables. Four versions of each “distanced” variation were created. Distancing condition was a between-subjects variable. The assignment of items to conditions was counterbalanced across four lists in each condition. The same 24 nonexperimental passages were used.

Procedure. The texts were presented, and reading times were collected, using MEL on 486 and Pentium PC computers. Otherwise the procedure was the same as in Experiment 1.

Results

After the experiment had been run, we found that two passages contained errors. One passage did not have inconsistent conclusions. The other was not distanced (i.e., not different from the Experiment 1 version). The passages are removed from the analyses. Mean reading times for the line containing the conclusion are presented in Table 6.

Filler between premises. When filler was inserted between the first and second premises (but the second premise was still adjacent to the conclusion), mean reading times were slower on the line containing the conclusion in the inconsistent than in the consistent condition; $F_1(1, 28) = 5.52$, $MSE = 178,126$, $F_2(1, 18) = 12.23$, $MSE = 71,284$. As can be seen in Table 6, the reading times for the inconsistent conclusion are much slower than those for the consistent conclusion when both premises are present. This interaction, the same pattern observed in Experiment 1, and in Experiment 2 when one premise immediately preceded the conclusion, was significant; $F_1(1, 28) = 8.29$, $MSE = 87,903$; $F_2(1, 18) = 21.37$, $MSE = 53,845$.

TABLE 6
Mean Reading Times (in ms) on Conclusions by Consistency and Number
of Premises by Distancing Condition in Experiment 3

	<i>Filler Between Premises</i>		
	<i>Consistent</i>	<i>Inconsistent</i>	<i>Difference</i>
Two Premises	2145	2471	326
One Premise	2197	2221	24
	<i>Filler After Premises</i>		
	<i>Consistent</i>	<i>Inconsistent</i>	<i>Difference</i>
Two Premises	2192	2271	79
One Premise	2262	2210	-52

Filler after premises. Neither main nor interaction effects approached significance by either the subjects or items test. Of primary interest is the interaction; $F_1 < 1$, $F_2(1, 18) = 1.44$, $MSE = 135,965$, $p = .25$. No differences were found for consistent versus inconsistent conclusions when both premises were present, or when only a single premise was presented. This also replicates the results from Experiment 2.

In summary, inconsistent conclusions were much less likely to be detected when the filler was placed after both premises and before the conclusion. Only when filler was placed between premises, with one premise still immediately preceding the conclusion, were inconsistent conclusions detected.

GENERAL DISCUSSION

The results of these studies indicate that readers are sensitive to inconsistent conclusions in scientific passages when all information is available, and when a conclusion appears immediately after the second premise. Inconsistent conclusions were not detected when readers were given only one premise, demonstrating that readers did not have much prior knowledge about these topics, and that both premises needed to be provided for the inference to be drawn. Also, inconsistent conclusions were detected only when the conclusion was immediately preceded by a premise. When even one sentence of filler information intervened between premise and conclusion, inconsistent conclusions were no longer detected. These results suggest that readers are able to generate causal bridging inferences only when they have knowledge of all necessary information, and when the necessary information is still accessible while they are reading the conclusion.

Our results are consistent with the findings of both Millis and Graesser (1994) and Singer and Gagnon (1999), both of which found evidence that causal bridging inferences can be computed by readers. In addition, there were several differences in methodology between our study and its predecessors that extend our knowledge of the conditions necessary for such inferences. Whereas Noordman et al. (1992) found evidence for causal inferences only when readers were given specific reading goals, and Singer and Gagnon (1999) found evidence for inferences using passages with explicit causal markers, we found evidence for inferences without either cue.

Another difference between our materials and those of Singer and Noordman, is that in these studies both premises were presented on their own lines, and were presented before a conclusion. This difference led to two novel findings. First, it extended previous studies by showing that causal inferences can be computed when information is presented in a deductive form. Since both premises were presented before the conclusion, the conclusion to be computed is actually a logical one. And, our results indicated that readers were able to recognize an incorrect logical

conclusion (the evidence for this is the longer reading times on inconsistent conclusion lines). Other research on the processing of logical conclusions from text has found similar results. Both Lea (1995) and more recently Rader and Sloutsky (2002) have found evidence that readers automatically compute the logical conclusions of “if P then Q” conditional forms embedded in discourse.

Second, because premises and conclusions were in separate lines, and passages were presented a line at a time, we were able to manipulate the availability of the premises at the point at which the reader encountered the conclusion. The result of this manipulation was that the contradiction effect was still obtained when one premise immediately preceded the conclusion, even though a line or two intervened between the two premises. However, when a line or two intervened between two adjacent premises and the conclusion, readers no longer showed the contradiction effect.

Although these experiments suggest when readers are able to compute inferences, neither these experiments nor others published previously provide information on exactly when causal inferences may be computed. Clearly the reader is recognizing when a conclusion is inconsistent with the potential causal inference. But, does the reader compute the inference spontaneously as soon as all necessary information is available? Or is the validation of the inference prompted by the target probe, the conclusion, or the contradiction? Longer self-paced reading times can result from either noticing the inconsistency with an already-computed inference or from computing the inference at the time when the conclusion is presented. The present data do not allow us to discriminate between these two alternatives. In order to determine if readers compute an inference as soon as they have read both premises, or whether an inference is computed only in a later stage of attempting to verify or integrate a conclusion, the collection of more detailed reading time data on both premises and conclusions is an important direction for future studies.

The results of these experiments indicate that causal bridging inferences can be drawn from scientific text when all the necessary information is available and accessible. The pattern of results also suggests that the availability of the potential inferences for instantiation or integration into the ongoing discourse model may be short-lived, as even one intervening sentence was enough to curtail the detection of inconsistent conclusions. This stands in contrast to results based on narrative texts in which inconsistent causal inferences can be detected even when separated from their antecedent information by several sentences or paragraphs (e.g., Klin, 1995). Causal inferences in scientific expository text do not appear to be computed or integrated in a spontaneous or durable manner as a necessary part of a coherence-building process. Instead, they seem to act more like elaborative or predictive inferences, which may be generated only partially or weakly, and instantiated or integrated only under specific conditions, or available for limited amounts of time (e.g., Cook et al., 2001).

Why do causal inferences appear to be spontaneously and durably encoded in narrative text comprehension and not expository? One possible factor is the structure of narrative text. There have been many demonstrations that the structure of narrative text is inherently causal, and that the memory representations of narrative texts are based on causal properties of the text (Trabasso & van den Broek, 1985). The structure of scientific text tends to be less uniform, less obvious, and less familiar to readers. Expository texts are frequently processed on a more descriptive level, and readers may be less motivated to make explanatory inferences from an expository text (Bock & Brewer, 1985; Goldman, 1997; Meyer, 1985; Miller, 1985). A related possibility is that readers may have different goals as they approach narrative and expository text.

Another likely factor is that differences in familiarity and knowledge between the content of narrative texts and scientific texts may govern which inferences can be made. In these studies, we attempted to give readers all the information they needed to make causal bridging inferences, and they were indeed successful at recognizing contradictions when the information that was explicitly mentioned was still available. However, even though we provided much of the critical information for each inference, the presence of background knowledge may have still affected the availability of distanced causal information. Even though all the information was given in our scientific texts, perhaps only readers with enough knowledge of the subject matter could access it once it had left working memory. Therefore, another reason readers may be able to make causal inferences even across long distances in narrative text is because of their expert knowledge about people's motives, intentions, and typical actions or event sequences, as opposed to their relatively poor knowledge of scientific topics.

This observation makes the importance of discourse markers in expository text more salient. Readers of expository text on unfamiliar topics apparently need the causal markers to remind them that premises were mentioned that could support the conclusion made in the text. Our results suggest that these markers may be necessary for readers to detect causal inconsistencies when the information needed for the conclusion is distanced. Indeed, a study by Singer and O'Connell (in press) has some suggestive results in this regard. While causal relations in narrative passages can be computed without conjunctions, such connectives are needed to guide the processing of readers of expository texts on unfamiliar topics. This highlights a practical need for writers of expository text to signal the argumentative or causal structure of their texts to the reader. Our studies suggest that without such signals, many connections will not be made.

These studies investigated when readers are able to generate and integrate causal bridging inferences into their representation of expository texts on scientific topics. The contribution of the present set of experiments is in demonstrating that intervening text can drastically affect whether inferences can be drawn from scientific text. When information is not immediately available, and when there is

also no marking of causal relations, nor explicit goals for reading, causal bridging inferences will not be made from scientific text.

ACKNOWLEDGMENTS

The authors thank Amy Brodhead, Sherrie Brown, Cara Jolly, and Ellen Southworth for their assistance with data collection; and Rob Mason, Keith Millis, and Murray Singer for their helpful comments on this research.

Portions of this research were carried out while Jennifer Wiley was a post-doctoral fellow at the University of Massachusetts at Amherst, supported by a training grant from the National Institutes of Mental Health (MH 16745). This research was also supported by a grant from the National Science Foundation to Jerome Myers (SBR 96-30740), and grants from the Office of Naval Research Cognitive and Neural Science and Technology Program (N00014-00-1-0264, N00014-01-10339) and the National Science Foundation (REC 0126265) to Jennifer Wiley. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these organizations.

REFERENCES

- Albrecht, J. E., & Myers, J. L. (1995). Role of context in accessing distant information during reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*(6), 1459–1468.
- Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*(5), 1061–107.
- Bock, J. K., & Brewer, W. F. (1985). Discourse structure and mental models. *New Directions for Child Development*, *27*, 55–75.
- Cook, A. E., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory & Language*, *44*(2), 220–234.
- Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. Van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169–193). Mahwah, NJ: Erlbaum.
- Duffy, S. A. (1986). Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 208–219.
- Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, *10*(6), 597–602.
- Goldman, S. R. (1997). Learning from text: Reflections on the past and suggestions for the future. *Discourse Processes*, *23*(3), 357–398.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning & Verbal Behavior*, *13*(5), 512–521.

- Just, M. A., & Carpenter, P. A. (1980). A Theory of Reading: From eye-fixations to comprehension. *Psychological Review*, 87, 329–354.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163–182.
- Klin, C. M. (1995). Causal inferences in reading: From immediate activation to long-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21(6), 1483–1494.
- Lea, R. B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21(6), 1469–1482.
- Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, 50(3), 643–655.
- McKoon, G., & Ratcliff, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning & Verbal Behavior*, 19(4), 369–386.
- McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Educational Psychology: Learning, Memory & Cognition*, 12, 82–91.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99(3), 440–466.
- McNamara, D., Kintsch, E., Songer, N., & Kintsch, W. (1996). Are good texts always better? Interaction of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1–42.
- Meyer, B. J. F. (1985). Prose analysis: Purposes, procedures and problems. In B. Britton & J. Black (Eds.), *Understanding expository text* (pp. 269–304). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Miller, J. R. (1985). A knowledge-based model of prose comprehension: Applications to expository text. In B. Britton & J. Black (Eds.) *Understanding expository text* (pp. 199–226). Hillsdale: Erlbaum.
- Millis, K. K., & Graesser, A. C. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory & Language*, 33(5), 583–599.
- Murray, J. D., Klin, C. M., & Myers, J. L. (1993). Forward inferences in narrative text. *Journal of Memory & Language*, 32(4), 464–473.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, 26(2–3), 131–157.
- Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20(4), 876–886.
- Noordman, L. G., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory & Language*, 31(5), 573–590.
- O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14(3), 410–420.
- Otero, J., & Kintsch, W. (1992). Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, 3(4), 229–235.
- Paris, S. G., & Lindauer, B. K. (1976). The role of inference in children's comprehension and memory for sentences. *Cognitive Psychology*, 8(2), 217–227.
- Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory & Language*, 27(4), 399–415.
- Rader, A. W., & Sloutsky, V. M. (2002). Processing of logically valid and logically invalid conditional inferences in discourse comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 28(1), 59–68.
- Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition*, 12(4), 315–328.

- Singer, M. (1988). Inferences in reading comprehension. In M. Daneman & G. E. Mackinnon et al. (Eds.), *Reading research: Advances in theory and practice*, Vol. 6. (pp. 177–219). San Diego, CA: Academic.
- Singer, M., & Gagnon, N. (1999). Detecting causal inconsistencies in scientific text. In S. R. Goldman, A. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 179–194). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Singer, M., Harkness, D., & Stewart, S. T. (1997). Constructing inferences in expository text comprehension. *Discourse Processes*, 24(2–3), 199–228.
- Singer, M., & O'Connell, G. (in press). Robust inference processes in expository text comprehension. *European Journal of Cognitive Psychology*.
- Suh, S. Y., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, think-aloud protocols and recognition priming. *Journal of Memory & Language*, 32, 279–300.
- Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory & Language*, 24(5), 612–630.
- van den Broek, P. (1990). Causal inferences and the comprehension of narrative text. In A. C. Graesser & G. H. Bower (Eds.), *Inferences and text comprehension* (pp. 175–96). San Diego, CA: Academic.
- Wiley, J., Mason, R. A., & Myers, J. L. (2001). Accessibility of potential referents following categorical anaphors. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27(5), 1238–1249.
- World Book. (1989). *Childcraft: The how and why library*. Chicago: World Book-Childcraft International.