Pictures and anaphora: Evidence for independent processes

ARTHUR M. GLENBERG and PETER KRULEY
University of Wisconsin, Madison, Wisconsin

Pictures enhance our comprehension of written texts, but the perceptual and cognitive processes that underlie this effect have not been identified. Because integrating the information contained in a text places demands on working memory, the effect of a picture may be to expand the functional capacity of working memory and thereby to facilitate comprehension. Reasoning thus, we predicted that the availability of a diagram would interact with the difficulty of resolving anaphoric references in texts. The resolution of an anaphor distant from its antecedent (which should stress working memory) should benefit greatly from the presentation of a picture, whereas the resolution of an anaphor near to its antecedent should benefit less from the presentation of a picture. Picture availability and distance separating the anaphor from its antecedent were manipulated in experiments involving both cumulative and moving window presentations of texts. Although picture presence and ease of anaphor resolution significantly improved comprehension of the material, no evidence was found for an interaction of these factors. The results are interpreted as consistent both with dual code theory and with aspects of working memory management that do not involve anaphor resolution.

It is a fact that pictures help people to learn from texts (see Willows & Houghton, 1987, for reviews of the literature). What is less certain, is exactly how pictures have this salutary effect. We designed the experiments reported here to investigate a possible beneficial effect of pictures on an important process of text comprehension—anaphor resolution. To foreshadow our results, we found, once again, that pictures enhance comprehension. Nonetheless, we were not able to adduce any evidence that the enhancement was due to a direct effect of pictures on anaphor resolution. We discuss several ways in which pictures can aid comprehension without affecting anaphor resolution.

Our tack was to investigate the general hypothesis that pictures enhance working memory management. Most theories of text comprehension (see, e.g., Kintsch, 1988) specify that an important aspect of comprehension is the building of a representation (see also Gernsbacher, Varner, & Faust, 1990), and that this representation is built with the use of the facilities of working memory. Because working memory is thought to be severely limited in the amount of information that can be represented at one time, the comprehender may find it necessary to shunt information into and out of working memory (that is, to manage it) while building the appropriate structures. Fletcher and Bloom (1988), for example, have investigated various strategies for controlling the contents of working memory.

ory during reading. Pictures may facilitate working memory management in several ways. For example, pictures can serve as an external memory, so that the reader can avoid searches of long-term memory or the text itself, when information is required. Larkin and Simon (1987) have demonstrated how pictures may enhance a search of this sort. Glenberg and Langston (1992) have demonstrated that pictures can lead to a reorganization of information in working memory. In that experiment, subjects read texts describing four-step procedures. The texts described both the content of the steps and the order in which the steps were to be performed. When reading without a picture, subjects tended to represent the steps in the order in which they were presented in the text. When the same texts were accompanied by pictures that illustrated the order of the steps when the procedure was executed, the subjects tended to represent the order of the steps as executed, rather than as described in the text.

Anaphor resolution is a comprehension process that is sensitive to working memory management. An anaphor is a word or phrase that is interpreted in relation to previous elements of the discourse (the antecedent or referent of the anaphor). Pronouns are prototypical anaphors, in that the interpretation of a pronoun (e.g., the word *it*) often requires reference to a previous description in the text. Nonetheless, anaphors may be noun phrases, synonyms of previous nouns, temporal referents such as *the former* (Jakimik & Glenberg, 1990), or even the absence of a part of the text (see Halliday & Hasan, 1976). Successful anaphor resolution is probably an important component in generating a sense of coherence, the sense that the sentences in a text belong together.

The process of anaphor resolution is sensitive to distance between the anaphor and its antecedent. It takes

longer to accomplish anaphor resolution when the anaphor and its antecedent are far from one another in the text than when they are close together (e.g, in adjacent sentences). Demonstrations of this distance effect may be found in Clark and Sengul (1979), Daneman and Carpenter (1980), and Ehrlich and Rayner (1983). A standard interpretation of this distance effect is that it reflects a type of working memory management. Antecedents that are near to their anaphors are likely to be represented in working memory (or foregrounded) when the anaphor is read. When the antecedent is far from the anaphor, the antecedent may not be in working memory, and it must be reinstated to resolve the anaphor. This reinstatement process takes time and produces the distance effect.

We anticipated that pictures would facilitate this reinstatement process and thereby provide for smooth and effective comprehension. Our reasoning was based on the idea that a picture may serve as an information source that is easier to search than either the physical text or the representation of the text in long-term memory (Larkin & Simon, 1987). This reasoning predicts an interaction between distance and whether or not a picture accompanies the text. When the antecedent and the anaphor are close in the text, the antecedent should be represented in working memory, so that the picture should make little difference. However, when the antecedent and the anaphor are far apart, the antecedent would need to be reinstated in working memory, and search of the pictorial representation should facilitate reinstatement.

Although we anticipated this interaction, our experiments were designed so that we could discriminate among several specific hypotheses. First, it might be that pictures need to be perceptually present to facilitate the reinstatement of antecedents (perceptual hypothesis).

Second, it might be that exposure to a picture creates a memory representation (of the picture) that can be used to facilitate reinstatement in the absence of the physical picture (memory hypothesis). Third, contrary to our expectations, pictures might not affect anaphor resolution, but might enhance comprehension through other means (general facilitation hypothesis).

The subjects read texts containing a short scientific description of a single object, typically a living organism, a part of an organism, or a mechanical device (some examples: the sphagnum sporophyte, the structure of a leaf, and a solar collector). Either four or five distinct parts of the object were described in the text. An antecedent sentence described the name and location of a part, and an anaphoric sentence repeated the spatial location, which was thus the anaphor, and presented one fact about the part. Each text conformed to the structure shown in Table 1. This structure includes two anaphoric sentences. One anaphoric sentence, the near anaphor, required recovery of an antecedent mentioned in the immediately preceding sentence of the text, Antecedent 2. The other anaphoric sentence, the far anaphor, referred to an antecedent mentioned three or more sentences previously, Antecedent 1. In both cases, the anaphor was a definite noun phrase describing a spatial location, such as, the part on the bottom. The antecedent and anaphoric sentences were short declarative sentences in which antecedent and anaphoric phrases occurred at the beginnings of the sentences. The filler and coherence sentence segments served two purposes. First, they increased the distance in the text between the far anaphor and its antecedent, and second, they improved the coherence of the texts. To control for effects of specific sentences on reading time and comprehension, the texts were designed so that either critical sentence could serve as a near or a far anaphor.

In addition, texts were sometimes accompanied by pictures. Figure 1 is the picture that could accompany the text in Table 1. Pictures showed the structural relationships between parts of the object in a level of detail typical of a high school or an undergraduate textbook. For many of the objects, the picture suggested a three-dimensional representation. Each part of the object named in the text was labeled in the picture.

On one third of the trials, a picture accompanied the text (full-picture condition). On one third of the trials, a picture accompanied the text, but it disappeared before the near anaphor sentence was displayed (disappearing picture condition). For the remaining texts, no picture was

Table 1
Sample Text Used in the Study

| Segment Name	Text
Introduction:	The sphagnum sporophyte is the reproductive system of a fungus that grows on plants.
Antecedent 1:	The operculum is at the top of the sphagnum sporophyte.
Fillers:	The fungus's spores are contained in an arc-shaped central column surrounded by a sterile region called the columella.
Antecedent 2:	The pseudopodium is on the bottom of the fungus.
Near anaphor:	The part on the bottom elevates the entire sporophyte.
Coherence Sentence 1:	Elevation increases the distribution of spores when they are released.
Far anaphor:	The part at the top maintains internal pressure.
Coherence Sentence 2:	This internal pressure is what will eventually cause the sphagnum sporophyte to pop and release the spores.
Near question:	What does the pseudopodium do? (a) releases spores. (b) elevates the sporophyte. (c) nourishes the sporophyte. (d) creates pressure. (Answer: b)
Far question:	What is the function of the operculum? (a) to elevate the sporophyte. (b) to distribute spores. (c) to maintain internal pressure. (d) to aid in attachment. (Answer: c)
Filler question:	Which part surrounds the central chamber? (a) foot. (b) columella. (c) pseudopodium. (d) operculum. (Answer: b)

The Sphagnum
Sporophyte
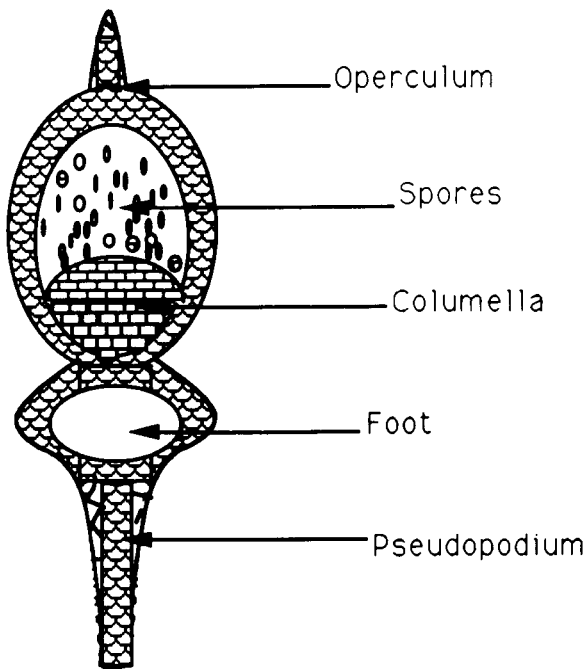
Figure 1. The diagram presented with the text of Table 1.

shown (no-picture condition). Consider the predictions from the three hypotheses outlined previously. First, suppose that the picture must be perceptually available to facilitate anaphor resolution. In this case, distance and picture condition should interact. That is, a distance effect should be found in the no-picture condition, but this distance effect should be substantially reduced in the full-picture condition. The disappearing picture condition should parallel the no-picture condition, because the hypothesis supposes that the picture must be perceptually present to facilitate the reinstatement of distant antecedents. Second, suppose that pictures result in a mental representation that can be used to facilitate the reinstatement of antecedents (memory hypothesis). In this case, there should again be an interaction between distance and picture condition, but the disappearing picture condition should parallel the full-picture condition (the distance effect should be reduced in both cases). Finally, consider the possibility that pictures enhance comprehension, but not anaphor resolution (general facilitation hypothesis). In this case, pictures will produce a main effect (better performance than in the no-picture condition), but there should not be an interaction with distance.

We used two dependent variables. One was reading time. The subjects pressed a key to expose successive parts of the text, and the time between keypresses was recorded. In this case, an interaction between distance and picture condition would be revealed by (1) an increase in read-

ing time with increases in distance in the no-picture condition, and (2) a smaller increase (or no increase) with increases in distance in the picture conditions. Unfortunately, this prediction may be compromised by the mechanics of the reading situation. That is, subjects must physically move their eyes from the text to the picture, and this will take some time. Thus, if subjects only refer to the picture in the far condition, the eye movement time will be added only in this condition, working against the predicted interaction. Also, having the picture present may induce other strategies, such as using the text to build a cognitive model of the object, and using the picture as a check for accuracy. Thus, the picture may slow reading, rather than speed it. For these reasons, we included performance on comprehension questions as a second dependent variable.

The near question required integration of information presented in the sentences containing the near antecedent and the near anaphor. That is, the Antecedent 2 sentence provided the name of a part (e.g., "the pseudopodium") and its location (e.g., "on the bottom"). The near anaphor sentence referred to the part by giving its spatial location (e.g., "the part on the bottom") and described the function of the part (e.g., "elevates the entire sporophyte"). The comprehension question required the association of the name (from the antecedent sentence) and the function (from the anaphoric sentence). Similarly, the far question required the integration of the information presented in the Antecedent 1 sentence and in the far anaphor sentence. When performance on the comprehension questions was the dependent variable, the interaction between distance and picture condition would be revealed by (1) a decrease in percent correct with increases in distance in the no-picture condition, and (2) a smaller decrease (or no decrease) in the picture conditions.

Before the subjects read the texts, we measured their reading spans. Daneman and Carpenter (1980) demonstrated that reading span was correlated with success in resolving distant referents of pronouns. Supposedly, readers with large reading spans can maintain potential referents more easily in working memory than can readers with smaller spans. On the assumption that pictures facilitate working memory management, we expected the following pattern of correlations. Memory span should correlate with performance in the far condition when there are no pictures and readers are forced to rely on their own memories. The correlation should be substantially reduced, however, when pictures are present and readers can use the pictures to supplement their own memories. Unfortunately, the correlations with reading span were all low and unsystematic across the experiments. Thus, we will not discuss the reading span measure further, except to indicate its presence in the procedure.

EXPERIMENT 1

Method

Subjects. The 48 subjects included both paid participants and students enrolled in introductory psychology courses at the University of Wisconsin at Madison. The latter received extra credit for

participating. All of the subjects were paid a cash bonus that was based on the total number of correct responses on the comprehension questions. Six additional subjects were replaced—5 because computer failures made it impossible to analyze their data, and 1 for failing to exceed chance performance on the comprehension questions.

Materials and Design. A total of 35 texts (3 practice, 32 experimental) were used, each with a corresponding labeled diagram (see the example in Figure 1). Comprehension questions, which were used to test the subjects for knowledge of both spatial and factual information, were designed to motivate the subjects to read all of the sentences carefully; as a further incentive, the subjects were informed in advance of the payment of 10 cents for each correct answer. Three multiple choice comprehension questions were prepared for each text (see Table 1). Correctly answering the near question required the subjects to integrate information found in the Antecedent 2 and near anaphor sentences. Similarly, a far question required integrating Antecedent 1 and the far anaphor. Filler questions were included to prevent subjects from identifying specific segments of the text on which to focus attention.

The texts and pictures were displayed side by side on a Macintosh color monitor. Upon responding to a prompt, the subject was shown the diagram and the title of the text. Each subsequent keypress revealed the next sector (see below) of the text. In the no-picture condition, the portion of the screen allocated for the picture remained blank, and only the text was shown; in the disappearing picture condition, the diagram disappeared after the keypress that revealed the near anaphor sentence (as identified in Table 1); in the full-picture condition, both the picture and the text remained visible until the keypress following the final word in the text. Thus the difference between the disappearing and full-picture conditions was the status of the picture while the subject was reading the critical anaphoric sentences.

Eight of the 32 experimental texts were used as unscored distractor texts. These texts, which were intended as a further guard against subject strategizing, included some comprehension questions that did not follow the near/far/filler scheme. In addition, for these texts, the picture disappeared after the presentation of a randomly selected word, rather than after the end of Antecedent 2. The same eight texts were used as distractors for all subjects.

Two factors were manipulated within subjects. One factor was picture presence: each subject read eight texts in each of the three conditions (no picture, disappearing picture, and full picture). The second factor was anaphor distance, with two levels: near and far. Each text contained a near and far anaphor, each with a related comprehension question. The order of the texts was randomized for each subject, with each text appearing equally often in the three picture conditions. In addition, each text was seen equally often in a standard order and with near and far anaphor sentences (and corresponding antecedent and coherence sentences) interchanged.

Procedure. The subjects were instructed to aim for comprehension. They were also instructed that pictures might disappear as if they had "turned the page of a book." Along with the display of the title and diagram, the text was represented by dashes corresponding to the letters in each word of the text. The reader controlled the rate of presentation by pressing the space bar when he or she was ready to read a new sector of the text. After the keypress, the next sector of the text became visible, replacing the dashes that had represented it. A sector of text was normally a single word, but two-word infinitive phrases and two-word phrases beginning with an article (e.g., "the operculum") were presented as single sectors. Previously exposed sectors remained visible on the screen, so that the display was cumulative. The reading time (time between keypresses, in milliseconds) on each sector of the text was recorded. For the practice trials only, after reading a text, the subjects answered the three multiple-choice comprehension questions by typing in the letter of the alphabet corresponding to the chosen answer. During the experiment proper, the subjects read a block of four texts and then answered the 12 questions relating to these texts. The questions were presented in groups of three, corresponding to the order of presentation of the texts. Within a group of three questions, the order of presentation was randomized. Immediate feedback was provided by reporting the number of correct answers on each text.

Results

After analyzing the data, it was found that 1 subject's total of 21 correct answers was more than three standard deviations below the overall mean; this subject was omitted from all analyses.

Reading times. Median reading times in each of the picture conditions for each subject were computed for the critical anaphoric sentences; the means of these median reading times are shown in Figure 2.[1] The expected ef-
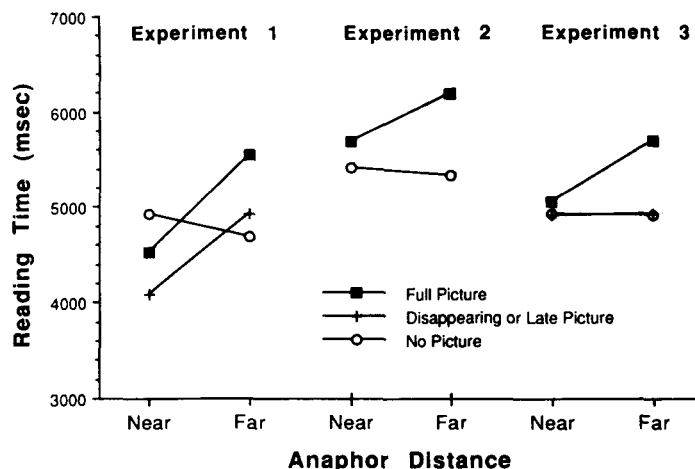

Figure 2. Mean of subjects' median reading times in Experiments 1-3, as a function of anaphor distance and picture presence.

fect of distance was found for the full- and disappearing-picture conditions. That is, subjects took longer to read the sentence containing the far anaphor than they did to read the sentences containing the near anaphors. There is also evidence of an interaction between picture condition and distance. Note, however, that this interaction is the opposite of the one predicted by the perceptual hypothesis and by the memory hypothesis. That is, instead of speeding reading by facilitating anaphor resolution in the far condition, the pictures appear to have slowed reading relative to the no-picture condition. These observations were confirmed by analyses of variance (ANOVAs) in which subjects (statistics reported as F_1) and texts (F_2) were the random factors. In all cases, an alpha level of .05 was used for the significance tests. The main effects of picture and distance and the picture × distance interaction were all significant [for picture, $F_1(2,92) = 3.27$, $MS_e = 2,036,189$, $F_2(2,46) = 6.67$, $MS_e = 975,448$; for distance, $F_1(1,46) = 10.78$, $MS_e = 1,875,759$, $F_2(1,23) = 31.29$, $MS_e = 532,208$; for the interaction, $F_1(2,92) = 5.98$, $MS_e = 1,856,294$, $F_2(2,46) = 7.62$, $MS_e = 820,365$]. Simple effects analyses indicated that the pattern of increasing reading time with increasing distance was significant in the disappearing picture condition [$F_1(1,46) = 7.26$, $MS_e = 2,280,009$, $F_2(1,23) = 15.00$, $MS_e = 626,394$] and the full-picture condition [$F_1(1,46) = 14.52$, $MS_e = 1,679,978$, $F_2(1,23) = 20.42$, $MS_e = 958,342$]; in the no-picture condition, in contrast, the simple effect of distance was not significant [$F_1(1,46) = .91$, $MS_e = 1,628,361$, $F_2(1,23) = .30$, $MS_e = 588,201$].

Comprehension. Mean accuracy of responding to comprehension questions was 81.1% (see Figure 3). Main effects were significant for both picture [$F_1(2,92) = 5.10$, $MS_e = 224.23$, $F_2(2,46) = 3.32$, $MS_e = 166.62$] and distance [$F_1(1,46) = 13.43$, $MS_e = 224.23$, $F_2(1,23) = $

5.90, $MS_e = 180.72$]. Predictably, accuracy increased as distance in the text decreased, and as picture availability increased. A central theoretical prediction involved a picture × distance interaction, but the interaction did not approach significance [$F_1(2,92) = 0.45$, $MS_e = 139.05$, $F_2(2,46) = 0.43$, $MS_e = 65.33$].

Analysis by reading speed. Individuals vary widely in reading speed, and this may reflect different reading strategies. To examine this possibility, the subjects were divided on the basis of overall reading time, and the data were analyzed using this factor. One subject with the median overall reading time was omitted from the analysis. No effects of reading speed on the comprehension measure were found (all F_1s < 1.15). For the sentence reading times, the three-way interaction of reading speed, distance, and picture was significant [$F_1(2,88) = 4.50$, $MS_e = 1,722,506$]. Simple effects analyses indicated the presence of a distance × picture interaction for slow readers [$F_1(2,88) = 10.21$], whereas this interaction was not significant for fast readers [$F_1(2,88) = .05$]. Figure 4 shows the patterns of reading times. A simple effect was calculated for the one case in which reading speed appeared to decrease with increasing distance—the slow readers in the no-picture condition—but this effect was not significant. Importantly, data from both the slow readers and the fast readers were contrary to the predictions derived from the memory hypothesis and the perceptual hypothesis, in that the presence of a picture never significantly speeded the reading of the far anaphor sentence.

Analysis by half of experiment. To rule out the possibility that theoretical factors might have been mediated by practice or fatigue effects, performance during the first and second halves of the experiment was analyzed in a three-way ANOVA. Interactions involving half were nonsignificant ($F_1 < 2.19$, $p > .10$ in all cases), and the
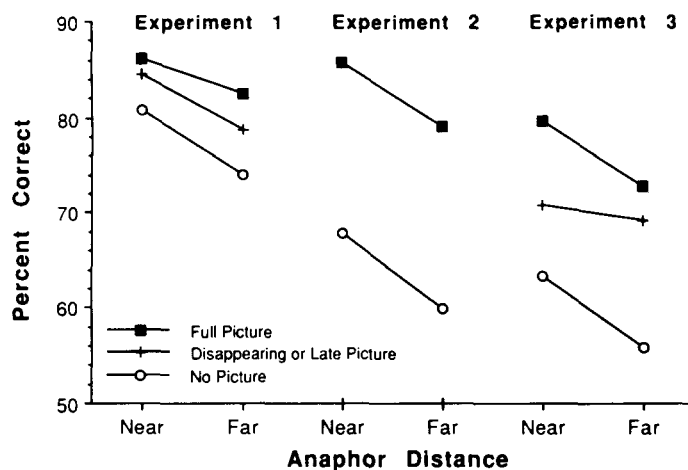

Figure 3. Mean percent correct in Experiments 1-3, as a function of distance and picture conditions. In Experiment 1, texts were presented cumulatively; in Experiments 2 and 3, a moving window presentation was used.
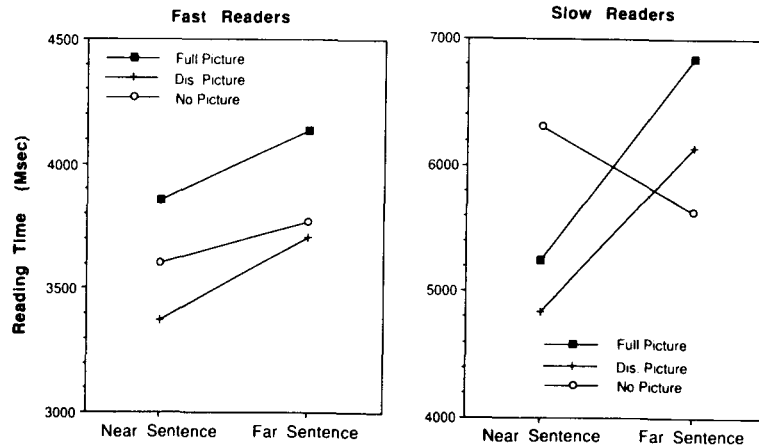
Figure 4. Reading times in Experiment 1, after a median split was performed on overall reading speed.

same pattern of results was obtained for Experiment 2, as well.

Discussion

We used distance as a conceptual marker for anaphor resolution, so that interpretation of the results requires the demonstration of a significant distance effect. Indeed, the distance effect was significant for both dependent variables; the resolution of far anaphors required more time and produced less accurate question answering than did the resolution of near anaphors for both the full and the disappearing picture conditions. No distance effect in reading times appeared for the no-picture condition, however. We do not know how to explain this absence. Perhaps the subjects found the task so difficult (in the absence of pictures) that they read all of the sentences very slowly. Analysis of the reading patterns of slow and fast readers separately revealed a more complex picture, but it did not change any important findings: where differences were significant, reading times increased with increasing distance and during reading with a picture.

Given that the distance manipulation was effective, what can we say about the three hypotheses described in the introduction? There was an interaction for reading times between the distance and picture conditions, but the form of this interaction was just the opposite of that predicted by the perceptual hypothesis and the memory hypothesis. That is, when subjects had pictures available, they read the far anaphor sentence slower than the near anaphor sentence. The interaction provides no support for the two hypotheses, but neither is the interaction fatal. As mentioned in the introduction, the picture may induce reading strategies that would countermand the predictions regarding reading times. Because of this possibility, we included the comprehension questions. Unfortunately for the perceptual and memory hypotheses, the comprehen-

sion question data do appear to be fatal. There was a distance effect, but little or no evidence for an interaction of distance and picture condition. Thus, the weight of the evidence is that pictures do not affect anaphor resolution.

EXPERIMENT 2

There are several reasons for being skeptical about the conclusion that pictures do not affect anaphor resolution. First, it depends on accepting the null hypothesis (we will present a power analysis in the General Discussion). Second, it rests in part on the comprehension question data, which are an off-line measure, whereas anaphor resolution is an on-line process. Third, it may be peculiar to the experimental context. That is, subjects may have become sensitive to the requirements of the comprehension questions and chosen to read exceptionally carefully, and perhaps reread the text, whether or not the picture was available. This strategy may have been encouraged by the continuous availability of the text. Experiment 2 was conceptually similar to Experiment 1, except for two changes. First, a moving window methodology (Just, Carpenter, & Woolley, 1982) was used. Each keypress exposed a single word (or short phrase) and erased the previously exposed text. This methodology eliminates the opportunity to reread the text, and thus should enhance the interpretability of the reading time data. Second, we dropped the disappearing picture condition for two reasons. It did not add much useful information in Experiment 1, and because subjects knew that pictures would disappear frequently, the subjects may have adopted unusual reading or picture inspection strategies.

Method

Subjects. Thirty-six undergraduate students at the University of Wisconsin at Madison contributed data. They received both course

credit for participating and a cash bonus that was based on the accuracy of their responses to the comprehension questions. Three subjects were replaced, 1 for failing to achieve performance above chance on the questions, and 2 because of computer failures.

Materials and Design. The texts and diagrams were the same as those used in Experiment 1, but 1 practice text and 4 distractor texts were omitted, leaving a total of 2 practice and 28 experimental texts, 24 of which were scored, for each subject. The one change from Experiment 1 in the presentation of texts and diagrams was that a moving window, as described by Just et al. (1982), was employed. As in Experiment 1, the initial display of the text consisted of dashes representing all nonspace characters in the text. Upon a keypress by the reader, one sector of the text appeared to replace the dashes corresponding to it. On the next keypress, the previous sector was replaced by dashes, and the next sector was revealed. Thus, the reading times measured represent the time during which the words in a sector of text were available for reading.

Picture availability and distance were again manipulated within subjects. Two levels of the picture variable were used: no picture and full picture. Therefore, each subject contributed 12 near and 12 far observations in each of the two picture conditions.

Results

Reading times. Because reading time for an individual word (or sector) is meaningful in the moving window presentation, analyses were performed on reading times for the anaphor phrases themselves as well as for the full sentences containing them. There were no significant effects on the phrase reading times ($Fs < 2.4, ps > .10$, for all analyses performed). Reading times for the full sentences are shown in Figure 2. The pattern of results is very similar to that found in Experiment 1. There was a significant effect of picture [$F_1(1,35) = 19.49, MS_e = 595,377, F_2(1,23) = 12.83, MS_e = 503,150$] and a significant picture × distance interaction [$F_1(1,35) = 5.28, MS_e = 590,623, F_2(1,23) = 6.13, MS_e = 257,488$]. The effect of distance approached significance for subjects [$F_1(1,35) = 3.54, MS_e = 484,865, .10 > p > .05$], but not for texts [$F_2(1,23) = 1.30, MS_e = 389,154$]. The same pattern of simple effects was found as in Experiment 1, with distance having a significant effect in the full-picture condition [$F_1(1,35) = 5.51, MS_e = 858,047, F_2(1,23) = 4.41, MS_e = 439,072$], but a nonsignificant

effect in the no-picture condition [$F_1(1,35) = .48, MS_e = 217,441, F_2(1,23) = .71, MS_e = 207,570$]. Note that this pattern is contrary to the predictions derived from memory and the perceptual hypotheses.

Comprehension. Percent correct responses to comprehension questions are shown in Figure 3. As in Experiment 1, main effects of picture [$F_1(1,35) = 89.7, MS_e = 136.10, F_2(1,23) = 26.03, MS_e = 312.53$] and distance [$F_1(1,35) = 25.66, MS_e = 74.73, F_2(1,23) = 12.71, MS_e = 100.46$] were both significant. The picture × distance interaction was not significant (both $Fs < 1$). Failure to find this interaction is contrary to both the memory and the perceptual hypotheses.

Analysis by reading speed. In Experiment 2, slow readers performed significantly better on the comprehension measure than did fast readers [80.7% vs. 65.5%; $F_1(1,34) = 11.57, MS_e = 715.55$]. For the sentence reading times, the three-way interaction of reading speed, distance, and picture was again significant [$F_1(1,34) = 5.89, MS_e = 558,655$]; for this experiment, the reading speed × distance interaction was also significant [$F_1(1,34) = 12.65, MS_e = 380,144$], and again simple effects analyses indicated the presence of a distance × picture interaction for slow readers [$F_1(1,34) = 14.21, MS_e = 558,655$], but not for fast readers [$F_1(2,88) = .11$]. As Figure 5 indicates, the basis of the interaction is that reading time increased with greater distance, but only for the slow reader group, and only when a picture was available. Analyzing the phrase time measure indicated that no interactions involving reading speed and other variables were significant ($F_1s < 2.70, p < .10$).

Discussion

Although the methodology used in Experiment 2 differs from that in Experiment 1 in several important respects, the results of the two experiments are quite similar. In particular, for the reading times, there was a significant interaction between distance and picture condition, but the interaction was just the opposite of that expected by the memory and perceptual hypotheses. Combined with the failure to find an interaction in the comprehension data, these are strong grounds for rejecting the hypotheses.
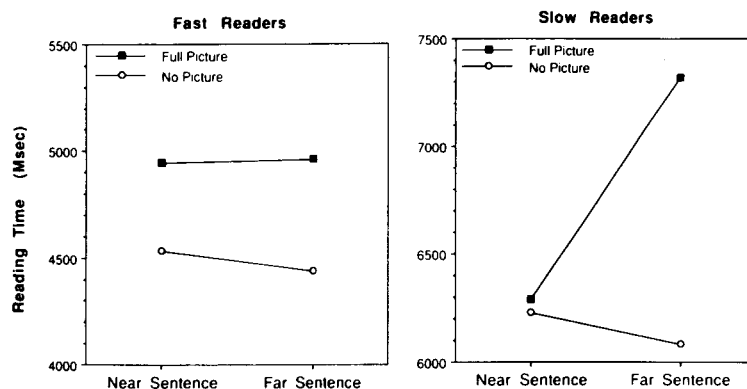

Figure 5. Reading times in Experiment 2, after a median split was performed on overall reading speed.

It is not surprising that the reading time measure revealed an effect of distance only for the full anaphoric sentences, rather than on the reading time for the anaphors themselves. Ehrlich and Rayner (1983) demonstrated just such an effect by using durations of eye fixations as the dependent variable. They concluded that only lexical access and syntactic parsing are carried out immediately upon fixation of a word; other processes, such as assigning a referent, are completed later. Van den Broek and Thurlow (1990) reported an analogous finding when measuring the time to make causal inferences. That is, the increase in reading time (with difficulty of the inference) occurred after the critical sentence demanding the inference, not during the reading of that sentence.

EXPERIMENT 3

The results of Experiment 2 replicated the failure to obtain an interaction between anaphor distance and picture presence; accordingly, we considered various hypotheses that might account for the obtained results. One such account is dual code theory (Paivio, 1986). In outline, processing the picture results in a long-term spatial representation, whereas processing the text produces a long-term verbal representation. The two representations are different in kind, but there may be associations relating concepts with the same referent. On this account, there is no reason to suspect that pictures would help in the on-line resolution of anaphora because that is a verbal task carried out separately from pictorial processing. Nonetheless, the pictorial representation may be useful when one is answering the comprehension questions. When the encoded verbal information is insufficient to provide an answer, the subject could consult the pictorial representation for the correct answer. There is no reason to suspect that the quality of the pictorial representation would vary with the linguistic variable of distance, and hence, the independence of distance and picture conditions is consistent with this approach.

A crucial difference between the dual code and working memory explanations is that only the latter postulates an on-line effect of picture presence. Consequently, the two accounts can be distinguished empirically: if the existence of a long-term pictorial representation is sufficient to enhance performance on the comprehension test, presenting subjects with a picture after they have read the text should be as facilitating as presenting a picture during the reading of the text. This distinction provided the rationale for Experiment 3, which included a late-picture condition, in which the picture was available for inspection only after the text had been read.

We considered three possibilities. First, pictures may have only an on-line effect. In this case, comprehension performance in the full-picture condition should exceed performance in the late-picture and no-picture conditions, and the latter two should not differ. Second, pictures may have only an off-line effect. That is, memory for the pictures is used to help answer the comprehension questions,

but the pictures do not facilitate on-line comprehension. In this case, performance in the full- and late-picture conditions should be comparable, and both should exceed performance in the no-picture condition. The third possibility is that both the working memory hypothesis and the dual code hypothesis are correct in that pictures affect on-line comprehension and that memory for the pictures can be used to answer questions after reading. In this case, performance in the full-picture condition (which has the benefit of on- and off-line facilitation) should exceed performance in the late-picture condition (which has the benefit of off-line processes only), and both of these should exceed performance in the no-picture condition.

Method

Subjects. Fifty-eight undergraduate volunteers contributed data; as before, they received course credit for participating, but no cash payments were made. One additional subject was replaced for failure to exceed chance performance.

Materials and Design. Three practice texts (the same as in Experiment 1) and 28 experimental texts (the same as in Experiment 2) were presented by using the moving window technique of Experiment 2. Anaphor distance (with two levels: near and far) and picture availability (no-picture, late-picture, and full-picture conditions) were manipulated within subjects. In the late-picture condition, the final keypress that dismissed the text display revealed the diagram of the object. The subjects were given unlimited time to study the diagram; they pressed the P key when they chose to dismiss the picture and continue to the next text. The subjects were not informed in advance whether a picture would follow a given text presentation.

Results

Reading times. The pattern of mean reading times for critical sentences, which resembles the results of the earlier experiments, is shown in Figure 2. The picture effect was significant in analyses by subjects and texts $[F_1(2,114) = 10.11, MS_e = 792,561, F_2(2,46) = 8.36, MS_e = 314,376]$. The effect of distance was also significant, with subjects as the random factor $[F_1(1,57) = 8.81, MS_e = 439,296, F_2(1,23) < 1]$. The interaction between picture condition and anaphor distance was significant by subjects, but not by texts $[F_1(2,114) = 3.84, MS_e = 967,932, F_2(2,46) = 2.65, MS_e = 215,919]$. Simple effects followed the same pattern as before: the effect of distance was significant in the full-picture condition $[F_1(1,57) = 10.31, MS_e = 1,095,877]$, but not significant in either of the other conditions $(Fs < 1)$. The three means obtained for reading the near sentence in the different picture conditions did not differ significantly: the simple effect of picture was nonsignificant at the near sentence $(F < 1)$, though significant at the far sentence $[F_1(2,114) = 10.62, MS_e = 1,066,601]$. Analyzing the reading times of the anaphoric phrases was generally uninformative.

Comprehension. The comprehension data, which followed the pattern of Experiments 1 and 2, are shown in Figure 3. Effects of picture $[F_1(2,114) = 29.05, MS_e = 277.203, F_2(2,46) = 30.69, MS_e = 107.963]$ and distance $[F_1(1,57) = 13.14, MS_e = 187.128, F_2(1,23) = 8.50, MS_e = 113.69]$ were significant; the interaction of

these factors was not [$F_1(2,114) = 1.79, MS_e = 178.077$, $F_2(2,46) = 1.39, MS_e = 95.022$]. Analyses done with the Tukey (a) method indicated that all pairwise differences between means of the picture conditions were significant. That is, comprehension was highest in the full-picture condition ($M = 76.1\%$), followed by the late- and no-picture conditions ($Ms = 69.9\%$ and 59.6%, respectively). This pattern is what one would expect if pictures both facilitated on-line comprehension processes and could be used when the comprehension questions are answered off-line.

Analysis by reading speed. Again, slow readers performed significantly better than fast readers on the comprehension measure [75.1% vs. 61.9%, $F_1(1,56) = 14.79, MS_e = 1,027.555$], but reading speed did not interact significantly with any other factor ($Fs < 2.28$). For the sentence reading times, the interaction of reading speed and distance was significant [$F(1,56) = 8.20$, $MS_e = 390,028$]. Simple effects analyses indicated that slow readers spent more time on the far than on the near anaphor sentence (6,460.0 vs. 6,057.4 msec), whereas reading times for the fast readers did not differ significantly (3,897.6 vs. 3,878.5).

Discussion

Because Experiments 2 and 3 differed in only one respect, it is not surprising that the principal findings of the previous experiments were replicated in Experiment 3. That is, far anaphor sentences were read more slowly than near anaphor sentences, especially by slow readers when a picture was available. On the basis of Experiment 3 alone, it appears that these reading time patterns can be more reliably generalized to new subjects than to new texts. The critical question being investigated in this experiment was the efficacy of a picture that was not available during on-line processing, and it was found that the late picture enhanced comprehension relative to reading with no picture, but was less effective than a picture that was available during reading. This finding may be seen as providing qualified support for both the dual code and working memory hypotheses; it will be examined in more detail in the general discussion that follows.

GENERAL DISCUSSION

In this discussion, we will consider two issues: possible problems with our conclusion that pictures do not facilitate anaphor resolution, and what accounts for the large facilitation in comprehension due to pictures.

There are several ways in which our results are limited, but not critically so. First, we used only two values of distance, near and far. Perhaps the values that we chose were not optimal for demonstrating effects consistent with the memory and perceptual hypotheses. This possibility can be dismissed readily. We did obtain distance effects in both reading time and performance on the comprehension questions. The problem was that we failed to find the predicted interactions. Also, Ehrlich and Rayner

(1983) manipulated distance over three values: near, intermediate (roughly corresponding to our near), and far. The largest difference that they observed occurred between the intermediate (our near) and the far conditions.

A second limitation is that we used only one type of anaphora, referring to a part by its spatial location. This is somewhat unusual in the literature, but it is just the sort of anaphor that should have revealed a benefit from a spatially organized picture. Nonetheless, no benefit (in anaphor resolution) was found. We used a different sort of anaphor—exact repetition of a part's name—in several other experiments. In some of those experiments, whole sentences were exposed with each keypress, and in some, the moving window technique was used. None of those experiments revealed any evidence in support of the memory or perceptual hypotheses.

Third, one could argue that repetition of a part's spatial location can be an inelegant form of reference. That is, when a concept is foregrounded (the current topic), it is easier to refer to the concept by using a pronoun instead of a noun phrase; full noun phrases seem to be required only when the concept is not the current topic (see Garrod & Sanford, 1990, and Gordon, Grosz, Gertner, & Rabin, 1990, for examples). Although this criticism is well founded, it predicts just the opposite of our results. That is, if the noun phrase is an inelegant form of reference in the near condition (because its antecedent is the subject of the previous sentence so may well be foregrounded), reading times should be slower in the near condition than in the far condition. A glance at Figure 2 will confirm that this did not occur.

Finally, our conclusion rests on accepting the null hypothesis of no interaction between picture condition and distance, and so the question arises as to whether or not the experiments had sufficient power to detect the interaction. In Experiment 1, there was sufficient power to have a 0.5 chance of detecting an interaction (gamma) as small as .284; in Experiment 2, an effect size of .325 could have been detected with a probability of 0.5; and in Experiment 3, an effect of .256 could have been detected with equal power. If the experiments are considered as independent replications, there was a greater than 0.875 probability of detecting an interaction effect as small as .325. Thus, it seems safe to conclude that the interaction is either absent or small, so that pictures are not affecting the resolution of anaphors.

There is no doubt that pictures facilitated comprehension (accuracy). The picture effect was significant in all three experiments (by subjects and by texts), and in Experiments 2 and 3 it was quite large; subjects averaged 62% correct in the no-picture condition and 79% correct in the full-picture condition. If pictures are not helping to resolve anaphors, what is behind this facilitation? The data are consistent with several possibilities, and inconsistent with at least one other. The one unlikely answer is that having a picture slows down reading and thereby produces better comprehension. Although the data are roughly consistent with this hypothesis (longer reading

times and better comprehension with pictures than without), a more finely grained analysis reveals trouble. For example, in the near condition in Experiment 2, there is little difference in reading time for full pictures and for no pictures, but there is a large difference in the comprehension scores. Also, in both experiments, subjects took longer to read the far sentence, but did more poorly on the far comprehension question.

One hypothesis that is consistent with much of the data is dual code theory (Paivio, 1986). The major benefit of the picture is to provide a mnemonic code that is useful in answering the comprehension questions, regardless of any on-line effect on comprehension. Clearly, the fact that performance in the late-picture condition exceeds performance in the no-picture condition (see Figure 3 and Experiment 3) is consistent with this hypothesis. Nonetheless, there is also evidence that the pictures facilitate on-line comprehension processes. First, having pictures available during reading did result in better performance than did examination of the pictures only after reading. Second, pictures did affect on-line behavior: the reading time for the far anaphor sentences was slower in the picture than in the no-picture conditions.

Given this evidence for an on-line effect of pictures, yet given also our strong evidence against the hypothesis that pictures facilitate the resolution of anaphors, what on-line processes are affected by the pictures? Consider the following two-part hypothesis. First, anaphor resolution is based on a search of verbatim and propositional representations of the text (see, e.g., McKoon & Ratcliff, 1980; O'Brien, Plewes, & Albrecht, 1990). That is, determining how one linguistic entity (e.g., *part at the top*) co-refers with another linguistic entity (*the operculum*), is a problem of syntax (both entities are singular) and semantics (an operculum can be a part). Solving the problem of co-reference does not require that the proposition being processed (e.g., operculum maintains internal pressure) be a true description of the state of the world. For example, the anaphoric sentence might have stated that *The part at the top is three miles long*. In this case, the anaphor, *the part at the top*, refers to the operculum, although the proposition that it is *three miles long* is blatantly false. Thus, at least the initial stage of anaphor resolution can be accomplished without regard to the actual state of the world as depicted in the picture.

The second part of the hypothesis is that once a new proposition is formed, however, the reader may attempt to understand how that proposition is an appropriate description of the situation. In this case, the picture, or a mental model based on the picture, may be helpful, because the picture is a representation of the situation. For example, how is it that the operculum maintains internal pressure? Examination of the picture reveals that the operculum is directly on top of the columella, and so it is in a position to exert a direct physical effect on the columella. Using the picture to encode new relations (e.g., that the operculum is physically adjacent to the columella) elaborates the text and enhances memory (Reder, 1979).

In short, anaphor resolution is not affected by the presence of pictures because it need not make use of a representation of a situation. Once anaphor resolution is accomplished and a proposition created, the picture is used to determine how the proposition is a correct description of the situation. This second process enhances memory for the information, and it is this enhanced memory that is reflected in the superior performance on the test in the with-picture condition.

On this account, why is reading slower with the far anaphor sentence than with the near anaphor sentence? Immediately prior to reading the near anaphor sentence, the subject has read the near antecedent sentence. In processing that sentence, the subject is likely to have looked at the picture and located the objects. Thus when the subject is determining how propositions from the near anaphor sentence apply to the situation, little time is required to locate where in the situation the new propositions apply. In contrast, new propositions from the far anaphor sentence occur in a context in which the objects have not been recently processed. In this case, it may take some time to locate where in the picture the new propositions apply.

This reasoning holds for processes directed at a picture that is physically present, as well as to a mental representation of the situation (a mental model) based on the picture. We have previously demonstrated that pictures facilitate the construction of limited-capacity spatial models (Glenberg & Langston, 1992). Thus, it is likely that subjects created mental models, based on the pictures, in the disappearing picture condition of Experiment 1. Note (in Figure 2) that for the disappearing picture condition, subjects spent more time reading the far anaphor sentences than the near anaphor sentences, even though the picture was not present while either sentence was being read. Why? Applying the reasoning from above, when one is checking the applicability of propositions from the near anaphor sentences, the relevant objects are likely to be represented in the mental model by virtue of having just read the near antecedent sentences. However, when one is checking the applicability of propositions from the far anaphor sentence, relevant parts of the mental model will have to be reinstated, taking additional time.

In previous work, we have made the claims that pictures help readers to construct mental models (Glenberg & Langston, 1992), and that mental models contribute to foregrounding and anaphor resolution (Glenberg, Meyer, & Lindem, 1987). Conjoining these two claims would seem to imply that pictures should facilitate anaphor resolution, but the data from Experiments 1-3 deny this. There is, however, a critical difference between the Glenberg et al. (1987) experiments and Experiments 1-3. In Glenberg et al., the stimuli were constructed to ensure that appropriate antecedents were represented in the mental model. As argued above, in Experiments 1-3, it is unlikely that appropriate antecedents were represented in the mental model, at least during reading of the far anaphor sentences. Thus pictures may well help to build mental models (Glenberg & Langston, 1992), and mental models

may well help to resolve anaphors (Glenberg et al., 1987); and yet, because of the limited capacity of the models, pictures will not facilitate anaphor resolution when the antecedents are not represented in the limited-capacity model at the time when the anaphor is processed (Experiments 1-3).

Although we have provided an account of our data, we are still somewhat puzzled. The pictures did not assist anaphor resolution, but there is no logical reason why they should not have done so. Even if initial attempts at resolution do not require a representation of the situation, when initial attempts fail, why didn't our subjects use the pictures? One possibility is that use of pictures while reading is a skill that must be learned. Whereas our subjects seem to have learned to use the pictures to aid in the construction of mental models (Glenberg & Langston, 1992) and to check on the interpretation of propositions (Experiments 1-3), they may not have learned to use pictures to aid in anaphor resolution. Direct instruction and training to use pictures in this way may produce an even larger enhancement of comprehension.

REFERENCES

CLARK, H. H., & SENGUL, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition, 7*, 35-41.

DANEMAN, M., & CARPENTER, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450-456.

EHRLICH, K., & RAYNER, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning & Verbal Behavior, 22*, 75-87.

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

GARROD, S., & SANFORD, A. (1990). Referential processing in reading: Focusing on roles and individuals. In D. A. Balota, G. B. Flores D'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 465-486). Hillsdale, NJ: Erlbaum.

GERNSBACHER, M. A., VARNER, K. R., & FAUST, M. E. (1990). Investigating general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 430-445.

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

GLENBERG, A. M., MEYER, M., & LINDEM, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory & Language, 26*, 69-83.

GORDON, P. C., GROSZ, B. J., GERTNER, A., & RABIN, K. (1990, November). *The role of pronominal reference in discourse segment coherence.* Paper presented at the meeting of the Psychonomic Society, New Orleans.

HALLIDAY, M. A. K., & HASAN, R. (1976). *Cohesion in English.* London: Longman.

JAKIMIK, J., & GLENBERG, A. M (1990). Verbal learning meets psycholinguistics. Modality effects in the comprehension of anaphora. *Journal of Memory & Language, 29*, 582-590.

JUST, M., CARPENTER, P., & WOOLLEY, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

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

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

McKOON, G., & RATCLIFF, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning & Verbal Behavior, 19*, 668-682

O'BRIEN, E. J., PLEWES, P. S., & ALBRECHT, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 241-249.

PAIVIO, A. (1986). *Mental representation: A dual coding approach.* New York: Oxford University Press.

REDER, L. M. (1979). The role of elaborations in memory for prose *Cognitive Psychology, 11*, 221-234.

VAN DEN BROEK, P., & THURLOW, R. (1990, November). *Reinstatements and elaborative inferences during the reading of narratives.* Paper presented at the meeting of the Psychonomic Society, New Orleans.

WILLOWS, D. M., & HOUGHTON, H. A. (1987). *The psychology of illustrations: I. Basic research.* New York: Springer-Verlag.

NOTE

1. The main analysis was done on the time to read the anaphoric sentences, rather than the time to read the anaphor sectors of the sentences, for two reasons. First, because of a programming error, the sector data were not available for all of the texts. Second, because the previously exposed text was available for inspection, some subjects might have elected to expose the whole anaphoric sentence (by multiple keypresses) before reading the sentence carefully. In any event, an analysis of the times spent on the anaphor sectors was in all essential respects like the analysis of the whole sentences. In addition, the methodology used in Experiments 2 and 3 will allow an analysis of the time to read the anaphor sectors as well as the time to read the anaphoric sentences.