JOURNAL OF MEMORY AND LANGUAGE 30, 385–406 (1991)

Locus of the Moses Illusion: Imperfect Encoding, Retrieval, or Match?

LYNNE M. REDER AND GAIL W. KUSBIT

Carnegie Mellon University

Five experiments are described that attempt to isolate the mechanism that produces the failure to notice discrepancies in questions or assertions, called the Moses Illusion (T. A. Erickson & M. E. Mattson, 1981, *Learning and Verbal Behavior*, 20, 540–552). Experiments 1 through 5 involved asking subjects to either (1) discriminate between distorted questions such as "How many animals of each kind did Moses take on the ark?" or (2) ignore the distortions and answer the question as if it were perfectly formed. Experiments 2, 3, and 5 also varied the familiarity of the assertions that were queried. Experiments 4 and 5 recorded reading times of the words in the question as well. Four alternative explanations for the Moses Illusion are considered: (1) the so-called illusion is just a cooperative response adopted by the listener/reader; (2) the retrieved memory structures are impoverished or incomplete and thus discrepancies cannot be detected; (3) the question is not carefully encoded and therefore the distorted word may not be fully processed during encoding; and (4) people often do incomplete matches between a complete representation of the question and a complete representation of the stored proposition that contains the answer. The evidence from these experiments does not support the first three alternatives, but is consistent with the partial-match hypothesis.

When asked *How many animals of each kind did Moses take on the ark?*, most people respond "two" even though they know that it was Noah, and not Moses, who took the animals on the ark (Erickson & Mattson, 1981). Only a few researchers (Bredart & Modolo, 1988; Erickson & Mattson, 1981; Reder & Cleeremans, 1990; van Oostendorp & deMul, 1990; van Oostendorp & Kok, 1990) have actually explored when and why these illusions occur, but the phenomenon seems quite robust. Indeed, a joke-question that school children ask is, "A plane crashed on the border between Mexico and the U.S. Where did they bury the survivors?" People often sense that the question is a joke, yet fail to detect the problem and still attempt to answer the question. There have been a large number of models concerned with sentence parsing and comprehension (e.g., Just & Carpenter, 1980; Kintsch & van Dijk, 1978; Schank, 1972) and a large number of models concerned with question answering (e.g., Camp, Lachman, & Lachman, 1980; Graesser & Murachver, 1985; Lehnert, 1977; Norman, 1973; Reder, 1982, 1987; Singer, 1984, 1985), yet none of these models addresses the issue of why people fail to notice this type of distortion. This paper is concerned with trying to understand the locus of the illusion. What are the mechanisms that underlie our failure to notice mismatches between what is actually presented to the comprehender and what the comprehender *believes* is being asked of him or her?

Erickson and Mattson (1981) found that

Send correspondence and reprint requests to Lynne M. Reder, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. The work reported here was sponsored by grants to the first author, National Science Foundation Grants BNS-003711 and IRI-8719469, and in part by the Office of Naval Research, Contract No. N00014-84-K-0063, Contract Authority Identification Number NR667-529. We thank C. Dennler, J. Martin, J. Siska, and C. Wible for help with various aspects of the experiments, and A. Miyake for comments on the manuscript.

0749-596X/91 $3.00

TABLE 1
SUBSET OF DISTORTED AND UNDISTORTED QUESTIONS USED IN EXPERIMENTS; ANSWERS AND ALTERNATE TARGETS

1. What kind of tree did Lincoln chop down?
2. By flying a kite, what did Franklin discover?
3. What did Goldie-Locks eat at the Three Little Pigs' house?
4. In which state did General Lee surrender to bring an end to the Civil War?

Question	Correct answer	Alternate target (distorted)	Alternate target (undistorted)	Correct answer to undistorted question
1	Can't say	—	Washington	Cherry
2	Electricity	Edison	—	Electricity
3	Can't say	—	Three Bears	Porridge
4	Virginia	Grant	—	Virginia

people frequently fail to notice a distortion when asked to answer a question or verify an assertion such as "Moses took two animals of each kind on the ark." The authors were careful to confirm that people who fell for the illusion did in fact know that it was Noah, not Moses, who took the animals on the ark. Even when people were warned that there may be some "tricky" or distorted questions, there was still a large tendency not to note the distortions until they were pointed out.

A common reaction to this result is that people are merely cooperating with the speaker. That is, people know what is meant by a question and therefore simply behave in a way that reflects the shared implicit knowledge. The first experiment reported here is intended to test the hypothesis that peoples' behavior is due to a tendency to be cooperative and that the Moses Illusion will not occur when subjects are explicitly instructed to watch out for it.[1]

If, on the other hand, people find it extremely difficult to monitor questions and sentences for distortions, then we need to ask why it is so difficult to notice certain substitutions of words. The paper will consider four explanations for the Moses Illusion, including the cooperative principle mentioned above. Specifically, is the failure to notice the mismatch due to imperfect encoding? That is, do people simply not "hear" the distorted word, or in the case of reading, do they simply "skip over" the critical word while reading? Another possibility is that the locus of the illusion is due to imperfect memory retrievals. That is, people may correctly encode the distorted question, yet their "fetch" into memory for the relevant information brings back only part of the required memory structure. The failure to notice the mismatch might arise because the retrieved proposition about the number of animals on the ark is missing the information about who was "commanding the ship" so to speak. Finally, it is possible that the illusion occurs because the matching process between the encoded question (or memory probe) and the retrieved information is imperfect. This would mean that the person answering the question encodes it adequately and also retrieves all the relevant information; however, the person fails to note an important mismatch between the retrieved information and the probe.

In the experiments of Erickson and Mattson, subjects were asked only a few questions, most of which were distorted. In the experiments to be reported here, subjects were asked to answer a large number of questions, half of them distorted. Table 1 presents a subset of the distorted and undistorted questions used in the experiments. The alternate words (either undistorted or distorted depending on the ver-

[1] A slightly different version of the Cooperative Principle is considered in the General Discussion.

sion of the question shown) and the answers to the questions are listed at the bottom of the table. Subjects were instructed that half of the questions would be distorted and that half would not be distorted. When subjects were assigned to the *Literal* task, they were to treat each question literally and not to give an answer if the question had been distorted; however, when subjects were given *Gist* instructions, they were told to ignore inappropriate words in the question and give answers to these questions as if they had not been distorted. The experiments to be reported here differed in a number of respects, but all shared these materials and had these basic task characteristics. To gain an impression of the difference between the Literal and Gist tasks, the interested reader might try the two tasks on the questions in Table 1.

Experiment 1

The motivation for the first experiment was to determine whether the Moses Illusion could be explained in terms of the Conversational Postulate (Grice, 1975). That is, conceivably people do notice the distortion and choose not to comment on it because they know what the speaker intended to say. If this position is correct, then it follows that subjects should find it easier (or at least no more difficult) to detect these distortions (Literal condition) than to ignore them (Gist condition).

Method

Materials. Examples of the questions used are listed in Table 1. The full set of questions appears in the Appendix. Sixty-two pairs of questions like the Noah–Moses question were constructed using the following constraints: (1) the term or phrase to be substituted in the question had to be semantically confusible with the original term; (2) it had to come from the same part of speech (syntactically); and (3) the distorted question could not be interpretable in a different manner such that there would exist a different, correct answer. For example, the following distorted question could not be used: *On what holiday do children go door to door, dressed in costumes, saying 'Ho, ho, ho'?* The intended answer is *Halloween,* but the expression "ho, ho, ho" is associated with Christmas, and therefore that question would produce a competing response to the intended answer "Halloween." Therefore, the distorted version of the Halloween question was written in the following form: *On what holiday do children go door to door, dressed in costume, giving out candy?*

It was also essential that the base form of the question be answerable in the absence of either the original or the substituted term. For example, "toll booth" was substituted for "phone booth" in the question *What comic book hero does Clark Kent become when he changes in a toll booth?* The substituted term invalidates the question but is irrelevant to figuring out the answer if the term is ignored. The substituted term did not always appear at the end of the sentence. Whether a distorted version of a question could create the illusion depends on the listener's knowledge about the topic and the two terms being substituted: knowing too much or too little makes it difficult to create the illusion. This issue will be addressed in the General Discussion.

Approximately half of the questions could be answered by reading only the information in the question that preceded the target word. The remaining questions required reading up through the target word or beyond it. We refer to the former type of questions as *Before* questions and to the latter as *After* questions. These two types did not differ in the number of words that they contained. Conceivably the pattern of responding in these tasks might depend on whether or not the answer to the question could be derived prior to reading the target word. For this reason, a number of the analyses to be reported (Experiments 2 and 3) include this Before vs. After factor.

The pairs of questions, one properly formed and one distorted, were essentially identical in length; however, the total set of questions used varied in length and, more importantly, in the number of content words associated with the answer. Some questions were rated as having only two words that are semantically related to the answer, and some were rated as having as many as six words associated with the answer. Three independent raters judged these questions for the number of associatively related content words, and there was almost no disagreement among them. Half of the questions had two or three related terms and the other half had four, five, or six related terms. The former half are called the *Few* terms condition, and the latter half are called the *Many* terms condition.

A pilot experiment was run on 15 subjects to screen our materials. A number of questions were replaced or eliminated after screening. Some were rejected because not enough people knew the answer, e.g., "What is the name of the woman who is opposed to ERA, adoption, and other liberal causes?" (Answer: Phyllis Schafly, substitute abortion for adoption.) Others had to be rejected because subjects often gave the wrong answer to either form of the question, e.g., "On which corner of the envelope should one put the return zip/(area) code?" The answer is "upper left"; however, subjects typically responded "upper right," assuming that the question concerned the location of the stamp.

Design and procedure. Subjects were randomly assigned to either the Literal condition or the Gist condition. The Literal group was asked to detect distorted questions, i.e., questions where one of the terms had been replaced with a related but inappropriate term, and answer "can't say" to these distorted questions. They were asked to give the correct answer to undistorted questions. In the Gist condition, subjects were asked to ignore distortions in questions and answer either version of a question as if there were no substitution of terms.

Two versions were made of each question, and subjects saw only one version of a given question. The assignment of questions to distorted versus normal was made randomly for each subject, with the constraint that half of the questions be distorted and half normal.

A subject was instructed about the nature of the task according to the assigned condition. Both groups of subjects were told that they would see one question at a time on the computer screen and to say their answer into the microphone as quickly as possible, while remaining accurate. In the Gist condition, subjects were told that:

> Some of the questions are "ill-formed" and if taken literally, do not have an answer; however, we want you to get the gist of the question and give the answer that you would give if the question were not ill-formed.

Subjects in the Literal condition were told:

> Some of the questions are "ill-formed" and if taken literally, do not have a correct answer. Please treat each question literally and answer "can't say" to the ill-formed questions.

Subjects controlled the presentation rate of questions by pressing a "next" button on a button box to view the next question. When the next question appeared on the screen, the clock started. When the subject spoke an answer, the voice key attached to a microphone stopped the clock. The computer program automatically recorded the answer time and erased the question from the screen. Both groups were instructed to say "don't know" if they did not know the answer. Subjects received implicit feedback in that the computer would display the answer that was expected, requiring the experimenter to type in an answer only if it differed from that one. For example, if the question was "What kind of tree did Lincoln chop down?" the answer "cherry" would be displayed on the screen for the Gist condition after a response, while the answer "can't answer" would be displayed on the screen for the Literal condition. Of course, if the question had used "Washington" rather than "Lincoln," both groups would see "cherry" as the answer.

The experimenter was present at all

times, both to type in the subject's response and to insure that the voice key was not triggered accidentally by an unintended vocalization. The experimenter also noted if the response time was valid due to a subject speaking too softly to trigger the voice key.

Subjects. Two populations of subjects participated in this study: one consisted of Carnegie Mellon alumni who had graduated less than 5 years earlier and volunteered to participate; the other group of subjects consisted of current Carnegie Mellon undergraduates who participated in order to help fulfill a course requirement.[2] There were 18 alumni in the Literal condition, 19 alumni in the Gist condition, 16 students in the Literal condition, and 14 students in the Gist condition.

Results

Approximately 4% of the trials were discarded because of inaccurate timing. Either the subject stopped the clock prematurely (e.g., by coughing into the microphone) or the microphone did not pick up the articulation and the clock did not stop when it should have. Of the remaining trials, analyses of variance were performed on both the correct RT data and the accuracy data using as factors answer condition (Gist vs. Literal), number of related terms (Few vs. Many terms), and question type (Normal vs. Distorted).

Table 2 presents the correct response time and accuracy data as a function of number of terms, task, and distortion. First, consider the response times. There is a sizeable and highly significant RT advantage for the Gist condition compared with the Literal condition, $F(1,63) = 6.1$ $p < .05$. Subjects were a tiny bit slower to respond to distorted questions, but this was not significant, $F < 1.0$. The number of words in the question that were associated with the answer also affected response time, $F(1,63) = 69.2$, $p < .001$, such that subjects were significantly slower when there were more associated terms. By itself, this result would not be interesting since this variable is confounded with sentence length; however, the accuracy data also show an effect of this variable.

TABLE 2
REACTION TIMES (S) AND ACCURACY AS A FUNCTION OF NUMBER OF TERMS ASSOCIATED WITH QUERIED INFORMATION, TASK, AND DISTORTED OR NOT: EXPERIMENT 1

	2–3		4–5–6	
	RT	% Corr	RT	% Corr
Literal				
Normal	3.77	76.28	4.73	79.39
Distorted	3.94	61.69	4.64	42.07
Gist				
Normal	3.34	84.28	4.05	77.79
Distorted	3.69	75.64	4.07	75.63

Almost all manipulations had an effect on the accuracy measure. First, subjects made significantly more errors when asked to interpret questions literally than when they processed questions in the Gist task, $F(1,63) = 22.0, p < .001$. This effect is due primarily to subjects in the condition where they are required to say "can't say," namely the distorted Literal condition. Subjects made significantly more errors in that condition than any other, resulting in a significant Question Type × Task interaction, $F(1,63) = 35.9, p < .001$. The fact that the error rate is so much larger for distorted questions in the Literal condition means that the difference in response times between the Literal and Gist tasks could be

[2] The former group served as the control condition in a study on aging. The latter group participated in a within-subject design such that they received the Literal task instructions for one block of trials, and the Gist task instructions for the other block. The assignment of task order was randomly determined for each subject. In order to use both populations of subjects, only their first block of trials will be considered. In this way, the results can be analyzed as if all subjects participated in a between-subject design. The excluded data from the second trial block was essentially indistinguishable from the first trial block.

viewed as an underestimate. That is, if subjects tried to be more careful in the Literal condition (so that accuracy would be as good for the distorted questions as for undistorted questions), responses would be slower still in the Literal condition.

Finally, the triple interaction of Question Type × Task × Number of related terms was also significant for accuracy, $F(1,63) = 22.6$, $p < .001$, such that the number of words in the question that are associated with the memory structure had no impact on performance except in the Literal task when the question was distorted. Subjects generally were least accurate with distorted questions in the Literal task (compared with undistorted questions or the Gist task), but the situation was exacerbated by having many terms in the question associated with the relevant memory structure. It may just be that it is harder to check every term in longer questions, but subjects seemed to have a bias to give the undistorted answer since their performance was unaffected by the number of terms if the question was undistorted.

One conclusion that seems clear from these data is that subjects found it easier to ignore distortions than to have to detect them. On the other hand, this research does not address the question of what types of distortions are difficult to detect and what types are easy. What is it about a sentence that makes it susceptible to the Moses Illusion? Consider the following statement taken from Bredart and Modolo's (1988) work on the Moses Illusion. It is a translation of a French statement that Belgian subjects had to verify or reject:

- *It was at the end of the 15th century that Magellan discovered America.*

It seems highly unlikely that any American subject would false alarm to a question that substitutes Magellan for Columbus. Perhaps it is the *strength* of the assertion being queried that affects one's susceptibility to the Moses Illusion: American and Belgian subjects differ in their respective knowledge concerning Columbus and Magellan. Most Americans know more about Columbus and probably less about Magellan than their Belgian counterparts. Conceivably for Belgians, the principal features associated with both Columbus and Magellan are that they were ocean explorers of the "New World" centuries ago.

The goal of the next experiment was to determine whether the Moses Illusion is due to an impoverished memory, i.e., whether strengthening the memory trace would affect tendency to fall for the illusion. For example, if Belgians knew the American mnemonic, "In 1492, Columbus sailed the ocean blue," they might not fall for the distortion. The result of Experiment 1, that the Moses Illusion was stronger with more terms related to the answer, suggests that amount of evidence for an interpretation (or activation relevant for an answer) influences processing, supporting the view that the illusion is memory-based. However, this result does not tell us the nature of memory involvement. Experiments 2 and 3 are intended to discriminate between the hypothesis that the locus of the illusion involves incomplete retrieval from memory and the hypothesis that it involves incomplete matching of the retrieved memory structure to the question, or simply incomplete encoding of the question.

EXPERIMENT 2

In this experiment, subjects were asked to study a series of facts prior to answering questions. These facts were a subset of those facts necessary to answer questions in the "Moses task." After the subjects had studied these facts, the remainder of the experiment was identical to the procedure of Experiment 1. The essential difference between Experiments 1 and 2 was that in Experiment 2, half of the questions had had the relevant knowledge made familiar before the questioning. Thus a subject might study the fact "Noah took two animals of each kind on the ark" before being asked to answer *How many animals of each kind did Moses take on the ark?* Regardless of whether the question would be assigned to the distorted or normal condition, the sentence that was studied always used the correct (undistorted) term. So, for example, a subject might later be asked *Who does Clark Kent become when he changes in a toll booth?* but the studied sentence would be *Clark Kent becomes Superman when he changes in a phone booth.*

Method

Materials and design. The materials, in terms of the questions, were exactly the same as those in the previously described experiment. The priming statements were simply the complete sentence answers to the undistorted form of a question. The priming statement (or answer) used as many identical clauses to the question as possible.

Half of the questions were randomly selected to be primed for each subject, with the constraint that half of the primed questions be from the distorted set and half from the undistorted set. Subjects always studied the undistorted statement during priming, which meant that the similarity of the priming statement to the question was a bit less in the distorted condition.

The other difference in design from Experiment 1 was that this experiment was a within-subject design, in which subjects received the two types of tasks in separate blocks. Half of the subjects were assigned to start with the Literal task, followed by the Gist task; the other half did the tasks in the reverse order.

Procedure. There was only one procedural change from the first experiment. In the first part of the experiment, subjects were told that they would be shown a series of statements to study. They were told to push the space bar to view the next statement only when confident that they had learned the information contained in the statement on the screen. In other words, study of the statements was self-paced. After studying the statements, subjects were given instructions for their first task.

Subjects. Twenty-one Carnegie Mellon undergraduates participated in the experiment to partially fulfill a psychology course requirement. Only native English speakers participated.

Results

Order of type of blocks for the two tasks (Gist vs. Literal) had no impact on performance. Therefore, order of testing is ignored in all reported analyses. A summary of the results of Experiment 2 is presented in Table 3. Mean correct response time and accuracy are reported as a function of task, whether or not the question is distorted, and whether or not the answer had been studied prior to the questioning phase. Because some of the condition means approached unity, an arcsine transformation was applied to the accuracy data before submitting them to an analysis of variance. The basic pattern matches what we found before, namely, that the Literal task is slower and less accurate than the Gist task, $F(1,19) = 28.9, 28.1$, respectively, $p < .001$. Distorted questions took significantly longer, $F(1,19) = 11.5, p < .01$, and were answered less accurately than undistorted questions, $F(1,19) = 17.7, p < .001$. The effect of distortion seemed to matter much more for the Literal task than for the Gist task with respect to accuracy, yielding a

TABLE 3
MEAN CORRECT RT (s) AND ACCURACY AS A FUNCTION OF TASK, DISTORTED/NORMAL, AND STUDIED/NOT STUDIED: EXPERIMENT 2

	Studied		Not studied	
	RT	% Corr	RT	% Corr
Literal				
Normal	3.51	93.54	4.65	78.19
Distorted	3.87	68.62	4.50	56.21
Gist				
Normal	2.47	96.85	3.26	72.11
Distorted	2.98	96.17	3.87	74.69

significant Task × Distortion interaction, $F(1,19) = 28.9$, $p < .001$. On the other hand, there was the suggestion of an interaction of Task and Distortion for response times, but with the effect being larger for the Gist task, $F(1,19) = 7.6$, $p < .05$.

Of particular interest is the effect on performance of priming, i.e., studying some of the statements. It is not surprising that there is a significant effect of priming on answer time such that previously studied statements are answered much faster than those that were not studied, $F(1,19) = 35.5$, $p < .001$. This may be due to an encoding advantage of having recently read the statement.

Subjects were also more accurate for studied statements, $F(1,19) = 75.5$, $p < .001$. The critical question is whether the improvement in accuracy can be attributed to an improvement in ease of detecting distortions. There was an improvement of over 12% in the Literal task for distorted questions; however, there was at least as large an improvement in all of the other primed conditions as well. In fact, the interaction of task, question type (Distorted vs. Normal), and priming was not at all significant, $F(1,19) < 1.0$.

The improvement in performance appears to be a knowledge effect where subjects become more familiar with the information and find it easier to answer correctly; still, subjects also improved when the correct information was irrelevant (should not be given—the Literal distorted condition). This suggests that familiarization might also enhance one's ability to detect distortions. Although we did not find clear evidence for enhanced ability to detect distortions with a stronger memory trace, we were not ready to abandon that hypothesis. It seemed to us that if one were asked a question about one's own mother, a distortion in the name would always be noted. So conceivably the manipulation was not strong enough: if the relevant information were really well learned, perhaps the Moses Illusion would disappear or at least be diminished. This was the motivation for Experiment 3.

Experiment 3

This experiment was a simple replication of Experiment 2 with only one modification: instead of studying the relevant information to their (self-determined) satisfaction, subjects were required to commit these assertions to memory. The purpose of this study was to ensure that the null effect in the previous study was not due to a manipulation that was too weak.

Method

Procedure. Subjects read through the priming statements carefully at their own rate as in Experiment 2. They were told to memorize them, but they were not told that the statements they were memorizing would be involved in a later experiment. After the memorizing phase, there was a recall phase during which subjects were required to recall the remainder of a studied sentence from the provided cue word or phrase from that sentence. If the sentence was not correctly recalled, the correct version would appear on the screen for the subject to study. The testing would continue for incorrectly recalled statements, using a single "dropout" procedure. The duration of this phase was 25 min on average.

After successfully completing the dropout recall phase, subjects participated in the same type of Moses Illusion questioning phase described for Experiment 2. The design and materials were exactly the same as before.

Subjects. Thirty undergraduates participated to satisfy part of a psychology course requirement. Only native English speakers participated. Subjects were randomly assigned to perform the Gist task first or the Literal task first.

Results

Once again, the order of performing the two tasks had no impact on performance and therefore this factor is ignored in all reported analysis. Response times and accuracy performance are displayed in Table 4 as a function of whether or not the relevant information had been memorized, whether the question was distorted, and whether the task was to treat questions literally or in a gist fashion. An arcsine transformation of the accuracy data was used in the analysis of variance. The basic pattern matches what we found before, namely, that the Literal task is slower and less accurate than the Gist task, $F(1,28) = 34.7$, 76.1, respectively, $p < .001$. Distorted questions were answered more slowly, $F(1,28) = 10.0$, $p < .01$, and less accurately, $F(1,28) = 100.1$, $p < .001$ as well. The interaction of Task with Distortion was only significant for accuracy, $F(1,28) = 36.6$, $p < .001$, indicating once more that subjects do exceptionally poorly in the Literal distorted condition.

Of more interest is the effect of priming on these answer times. It is not surprising that there is a significant effect of priming such that memorized statements are answered much faster than those that were not memorized, $F(1,28) = 49.0$, $p < .001$. Again, this may be due to an encoding advantage of having recently read the statement. The biggest effect of priming comes in the accuracy data: subjects were much more accurate for memorized statements, $F(1,28) = 122.9$, $p < .001$. By having just memorized the critical fact, subjects obviously knew the answer to that question.

TABLE 4
MEAN CORRECT RT (s) AND ACCURACY AS A FUNCTION OF TASK, DISTORTED/NORMAL, AND MEMORIZED/NOT MEMORIZED: EXPERIMENT 3

	Memorized		Not memorized	
	RT	% Corr	RT	% Corr
Literal				
Normal	3.24	98.75	4.47	68.11
Distorted	3.49	59.03	4.49	35.52
Gist				
Normal	2.38	98.13	3.33	79.72
Distorted	2.79	92.50	3.66	76.49

The most interesting result of all is the fact that priming, the familiarization variable, did not interact with the other variables of interest. Neither task nor distortion produced significant interactions with priming. This result is discussed further below.

Discussion

Contrary to expectations, the Moses Illusion is not diminished by making the relevant knowledge very familiar. In both Experiments 2 and 3 the familiarization variable appears to be *additive* with the other variables that produce the Moses Illusion. Priming resulted in a full 1-s savings in response time, yet this manipulation did not make it easier to detect a mismatch between the underlying information and a distorted question. If priming had affected the matching process (e.g., by bringing a more complete memory representation into working memory), then the improvement in response time due to this variable should have been greater for the Literal distorted condition than for the Literal undistorted condition. If familiarity affected one's ability to detect mismatches, then the Gist condition would be expected to suffer relative to the Literal condition; that is, distortions should be more salient when the knowledge has been primed, meaning that it would make it easier to notice distortions and harder to ignore them. None of this occurred.

The *additivity* of the priming or familiarization effect can also be seen in the accuracy data. There is an interaction of task with distortion on accuracy, such that distorted questions hurt performance more in the Literal task; however, this is true regardless of whether or not the question was previously studied. For accuracy as well as response time, studying the information improved performance but did not differentially improve detection of distortions.

Error Analysis

This conclusion is supported by another analysis. We looked at the types of errors

TABLE 5
MEAN ERROR RATES BY TYPE: EXPERIMENT 1

	Literal task				
	Undistorted answer	"Can't say"	Don't know	Wrong answer	Total
Normal	—	.11	.14	.07	.32
Distorted	.33	—	.13	.04	.50

	Gist task		
	Don't know	Wrong answer	Total
Normal	.15	.08	.23
Distorted	.20	.10	.30

Note. These data are only from the college students. For technical reasons it was impossible to reanalyze the data from the alumni.

that were made in the various conditions in order to distinguish between errors due to lack of knowledge and errors due to incomplete matching between the input and the memory structure. There were four types of errors when performing the Literal task: a subject could say "don't know," give the undistorted answer when the question was distorted, say "can't say" when the question was not distorted, or give an unrelated, wrong answer. In the Gist condition, errors fell into only two categories, *don't know* and *wrong answer*.

It is informative to compare types of errors in an experiment that did not involve a familiarity manipulation, e.g., Experiment 1, with one that did. First consider the data from Experiment 1 in Table 5. Notice that the proportion of errors for "wrong answer" and "don't know" is essentially the same for both the Literal task and the Gist task regardless of whether the statement was distorted or not. This confirms the presumption that the larger error rate in the Literal task (compared with the Gist task) is due to giving the undistorted answer to distorted questions and to occasionally saying "can't say" to a question that was not distorted.

Table 6 presents the same analysis for Experiments 2 and 3 combined, except that now the data are also partitioned into primed and unprimed categories.[3] For the unprimed questions the data look very similar to those in Table 5. Subjects did not know the answer (either wrong or "don't know" responses) to about 20% of the queries. These data are in stark contrast to those from the primed conditions. In the primed conditions, subjects say "don't know" to less than 3% of the questions. This shift is reasonable since they have just been studying the answers. On the other hand, studying the answers does not improve one's ability to detect distortions: subjects give nearly as many undistorted answers to distorted questions in the primed condition as in the unprimed conditions.

Earlier in this paper we mentioned that the stimulus questions could be categorized (post hoc) into two types, namely, those that can be answered *before* the target word is read and those than can only be answered *after* the target word is read. Conceivably the tendency to give the undistorted answer in the Literal distorted condition occurs be-

[3] The same analysis was also done separately for Experiments 2 and 3. The patterns were the same, with the same range of error rates and a correlation of .95.

TABLE 6
MEAN ERROR RATES BY TYPE: EXPERIMENTS 2 AND 3

	Literal task				
	Undistorted answer	"Can't say"	Don't know	Wrong answer	Total
Normal					
Unprimed	—	.07	.16	.05	.29
Primed	—	.03	.01	.01	.05
Distorted					
Unprimed	.35	—	.17	.03	.55
Primed	.31	—	.02	.01	.34

	Gist task		
	Don't know	Wrong answer	Total
Normal			
Unprimed	.16	.08	.24
Primed	.01	.02	.03
Distorted			
Unprimed	.18	.08	.26
Primed	.02	.04	.06

cause the reader has already received enough information to respond correctly before reading the distorted target word. This hypothesis can be evaluated by analyzing the size of the illusion as a function of whether or not the question is a *Before* or *After* question: When the distorted term comes late in the question, and the subject can already figure out the answer (can answer *Before* seeing the distorted target), will the illusion be larger than when the subject cannot figure out the answer until *After* the distortion? We consider the most relevant statistical comparisons to be (a) whether the Before versus After variable interacts with the effects of interest, and (2) whether the effects described previously hold for the *After* questions.

To address these questions, we performed an analysis of variance on the Literal task data for Experiments 2 and 3 combined, adding the Before versus After factor. There were no significant effects associated with the Before versus After variable for accuracy. For correct response times, however, subjects were significantly slower to answer a question if the question could be answered before seeing the target (4.15 vs. 3.90 s). This effect of slower response times for Before questions is due primarily to the distorted version of the target (4.4 for Before distorted; 3.9 for After distorted and for both types of undistorted questions). Indeed, there is a significant interaction of Before versus After with the type of question (Distorted vs. Normal), $F(1,49) = 7.0$, $p < .01$. Given that there was no difference in the propensity to fall for the illusion for the two groups of questions, we are uncertain how to interpret the difference in response times. It would seem that if subjects were answering the question without processing the distorted word (in the Before condition), then they should be *faster* than in the After condition; however, they are slower. This seems inconsistent with the suggestion that the effect is due to answering without attending to the question in its entirety. Questions should be read faster if they are not completely processed. Of course, those incompletely processed questions would be excluded from the response times, so the relevant comparison has to be error rates. The error rate data did not differ as a function of question classification.

To try to address the question of whether the Moses Illusion was due only to the Before questions, we performed an ANOVA on the After questions from Experiments 2 and 3. All of the main effects and interactions that were significant for the entire set of questions are significant for this subset. For example, in this analysis, accuracy for the Gist questions and the undistorted Literal questions averaged 83% but was only 57% for the distorted Literal questions, yielding a significant Task × Distortion interaction, $F(1,49) = 19.35, p < .01$. There was also a significant interaction for these two variables on RT, $F(1,49) = 22.95, p < .01$. Therefore, we believe that we can reject the explanation that the illusion is due to retrieving the answer from memory before processing the part of the question that contains the distortion.

Given the general conclusion from Experiments 2 and 3, that the basic Moses Illusion is essentially unaffected by whether the information had been primed and made more accessible, it seems clear that *imperfect memory retrieval* is not the cause of the illusion. If the illusion were caused by failing to retrieve the *complete* memory structure back into working memory, then one would expect a decrement in the illusion along with the improvement in answerability of the questions when the relevant knowledge traces are strengthened. At the beginning of this paper we considered four candidates for the illusion, including a "cooperative" principle that stipulated that the effect was due to people's habit of "overlooking" the errors of others. Experiment 1 made that explanation seem untenable unless one posits some kind of automaticity of cooperative processing, i.e., non-stopability. Even an "automatic" cooperative principle has difficulty because not all slips or distortions are treated equally (e.g., "Nixon" is noticed in the Moses question).

The data of Experiments 2 and 3 have cast doubt on the explanation of imperfect memory retrieval, leaving two plausible candidates for the locus of the illusion: the encoding of the question or memory matching. That is, it may be that the reason the illusion occurs is that people do not carefully listen to or read the questions and simply assume that they know already what the questioner is going to ask once they hear part of the question. The other possibility is that the answerer adequately encodes the question and adequately retrieves relevant knowledge from memory, but merely does a "sloppy" job of matching the question probe to the retrieved (from memory) data structure.

The next experiment attempts to determine whether the illusion is caused by a failure to carefully encode the question. It uses reading times of individual words of the question to try to address this issue.

Experiment 4

The logic behind this study is as follows: If the Moses Illusion is caused by a failure to carefully encode some of the words or the question, then the pattern of reading times for words of the question should vary depending on whether or not an error was made. For example, consider the situation where a subject errs to a question in the Literal distorted condition by giving the undistorted answer. If the illusion is caused by failure to carefully encode, then one would expect that the reading time for a *distorted target* word in the Literal condition would be less when the subject failed to notice the distortion than when the distortion was noted (and responded "can't say").

Method

Design and procedure. The procedure was essentially the same as Experiment 1, except that the subject controlled the onset of each word, and only one word of the sentence was displayed at a time. The program that we employed was written to be similar to the *moving window* paradigm developed by Just, Carpenter, and Woolley (1982). The sentence was hidden, except for the one word that the subject was reading. The subject forced the "window" along from left to right by pressing the space bar whenever he or she was ready to read the next word. Just et al. have argued that the word level effects from this paradigm are similar to those obtained from eye-fixation protocols collected from readers of normal text.

Subjects read one word of the question at a time except that compound words (e.g., New York) appeared together on the screen. Subjects knew when the question was completed because it ended with a question mark. At that point, they gave their answer into a microphone. The time to read each word was recorded as well as the time to give the answer (from the onset of the question mark).

The other difference in design from Experiment 1 was that this experiment was a within-subject design, in which subjects received the two types of tasks in separate blocks. The order of the task block was counterbalanced across subjects.

Subjects. Twenty-three college-age subjects were recruited from ads on the Carnegie Mellon and University of Pittsburgh campuses. They received $6.00 for participating in the 1-h experiment. Only native English speakers participated.

Results and Discussion

Target word reading times and accuracy data are displayed in Table 7. The accuracy data are comparable to those from the other studies that did not use the moving window paradigm and are not of primary interest. For the most part, the analyses will focus on word reading time data as a function of condition, accuracy, and type of error.

First we analyzed word reading times as a function of *type of word* and whether the question was answered appropriately. We considered three types of words: pretarget content words, target words, and post-target content words. (Function words,

TABLE 7

EXPERIMENT 4: TARGET READING TIMES (ms); PROPORTION OF OBSERVATIONS FOR CORRECT AND INCORRECT RESPONSES (IN PARENTHESES) (ERROR DATA ALSO BROKEN DOWN BY TYPE)

	Literal task					
			Errors by type			
	Correct	Errors	Undist	Can't say	Don't know	Wrong ans
Normal	525	515	—	559	535	546
	(.79)	(.21)		(.08)	(.10)	(.03)
Distorted	539	633	628	—	650	1002
	(.57)	(.43)	(.32)		(.09)	(.02)

	Gist task			
			Errors by type	
	Correct	Errors	Don't know	Wrong ans
Normal	429	618	600	426
	(.82)	(.18)	(.11)	(.07)
Distorted	441	771	832	747
	(.76)	(.24)	(.14)	(.10)

read more rapidly, were ignored.) A target word was defined as either the distorted word or the word that would have been replaced had the question been distorted. An appropriate or correct answer was defined as in previous experiments: giving the intended answer to the question regardless of distortion when in the Gist condition; answering "can't say" to distorted questions when in the Literal condition. For correct responses, reading times were 435 ms for pretarget content words, 484 ms for target words, and 509 ms for post-target words. For incorrect responses, reading times were 452 ms for pretarget content words, 628 ms for targets, and 594 ms for post-target content words.

Subjects read words more slowly on trials where they made errors, $F(1,21) = 8.7$, $p < .01$. Overall, targets were read significantly more slowly than content words, $F(1,21) = 11.8$, $p < .01$; the difference between them was larger on error trials, yielding a significant Type of Word × Accuracy interaction on word reading times, $F(1,21) = 6.2$, $p < .05$. Since targets are the critical words in the sentence and the ones for which the differences seem largest, most of the remaining analyses will focus on reading times to the target words.

Table 7 presents reading times for target words partitioned by task (Gist vs. Literal), type of question (Distorted vs. Normal), and type of answer (correct or type of error). There was no main effect of task on target reading time: subjects were faster to read words in the Gist task than in the Literal task when correct, but slower when they answered incorrectly. This interaction of task and accuracy on target reading time reached significance, $F(1,21) = 4.6$, $p < .05$.

It is interesting to note that the slow target reading times for both tasks, Gist and Literal, seem to occur when the target is distorted and an error is made. When subjects answer correctly, there is little evidence of a slowdown due to the distorted term.

The most important comparison involves the reading times for distorted targets in the Literal task, comparing times when the undistorted answer is given and when the subject correctly answers "can't say." Contrary to expectations, subjects read distorted targets as quickly as undistorted targets when the question was answered correctly. Furthermore, reading times for distorted words were *not* faster when they failed to notice the distortion than when the distortion was noticed. If anything, subjects read the distorted word slower when the distortion was not noticed: 628 versus 539 ms.[4] Thus, there does not seem to be evidence to support the view that the Moses Illusion is caused by inadequately encoding the distorted term: If reading time for critical words is any indication of the amount of encoding or processing time spent on that word, subjects should have read the distorted word faster when the illusion occurred than when they noted the distortion.

Experiment 5

The purpose of Experiment 5 was to look at the effect on reading time of familiarizing subjects with the queried information. If subjects are familiarized with the queried information prior to being asked to answer it, we would expect to replicate the knowledge advantage shown in Experiment 2, and additionally to show a reading time ad-

[4] This difference was reliable by a sign test (of the non-empty cells and non-ties, 26 out of 38 comparisons were in the direction of longer reading times for targets when an error was made), $p < .05$. The only explanation we can think of is an item-by-subject selection artifact. Subjects are faster to read those questions that they know well and those are the ones that they find easy to detect. This explanation has problems as does one postulating a general slowing on error trials (e.g., becoming dull-witted on certain questions): both would predict the same pattern for content words. However, this same contrast for content words was not significant by the same test (of the possible comparisons, 16 out of 37 were slower on error trials).

vantage for questions that were highly similar to the sentences that had been studied. The reading time advantage should occur because of physical or lexical priming; however, a reading time advantage should also occur because of the recent activation of the relevant knowledge structures. In this situation, one might actually expect subjects to be differentially slowed to the distorted target word. First, unlike the rest of the primed question, the distorted target word would not be studied during priming. Second, given that the relevant sentence is now well known, the violation of expectancy should be especially large. This experiment was designed to ascertain whether subjects would be appreciably slowed by the distorted target when the question had been primed, and if so, whether there was a relation between the tendency to read slower and the tendency to detect the illusion.

Method

The materials were the same as those used in the previous experiments; namely, the same questions and the same familiarization sentences were used. The familiarization procedure was identical to that used in Experiment 2. Reading times were collected in the same manner as Experiment 4.

Subjects. Twenty college-age subjects who read ads on the Carnegie Mellon and University of Pittsburgh campuses were recruited and received $6.00 for participating in the 1-h experiment. Only native English speakers participated.

Results

Target reading times are displayed in Table 8 as a function of whether the question was answered correctly, task (Literal vs. Gist), whether the question was distorted or not, and whether the question had been primed (recently studied statement similar to the question). Not surprisingly, reading times are longer when subjects erred than when they answered correctly, $F(1,18) = 14.1$, $p < .01$. The only exception is the primed statements in the Literal task when the question was distorted. In that case, subjects read the distorted word somewhat faster (although not significantly) on error trials.

It is interesting to note that although priming produced much faster reading times overall in both this experiment and in Experiment 2, reading times for target words in the Literal condition were slower when the question had been primed. In fact, the interaction of task and priming was significant, $F(1,18) = 16.2$, $p < .01$.

Of special interest are the target reading times as a function of priming and type of error made. Table 8 also displays reading times and proportion of observations per type of response for the Literal task. The unprimed data show the same pattern as in Experiment 4; viz., subjects read the distorted term slower when they fail to note the distortion (539 ms) than when they do notice the distortion (476 ms). The familiarization manipulation clearly had an effect on accuracy, raising it by about 20% correct. On the other hand, as in past experiments, this priming did not appreciably lower people's tendency to fall for the Moses Illusion. The reading times of special interest are for the distorted targets when the question had been primed in the Literal task, comparing the situation when the correct answer is given to when the undistorted answer is given. In this case, subjects are slightly faster when they fail to notice the distortion, 545 versus 556 ms.

Of course this difference is not statistically reliable, but it is interesting to note that if the information on which the question is asked is made more available (i.e., priming), the distorted term becomes more salient and takes longer to read than when the question had not been so primed. The improvement in detection, however, seems to be only about 5%, comparable to the improvement in Experiments 2 and 3.

In the Gist task, however, where subjects are not asked to monitor for distortion, subjects do not read the distorted term any

TABLE 8
TARGET READING TIMES (ms) AND PROPORTION OF OBSERVATIONS FOR CORRECT AND INCORRECT RESPONSES (ERROR DATA ALSO BROKEN DOWN BY TYPE): EXPERIMENT 5

	Literal task					
			Errors by type			
	Correct	Errors	Undist	Can't say	Don't know	Wrong ans
Normal						
Unprimed	453	593	—	773	598	586
	(.75)	(.25)		(.05)	(.15)	(.05)
Primed	459	840	—	407	839	748
	(.90)	(.10)		(.03)	(.05)	(.02)
Distorted						
Unprimed	476	563	539	—	642	450
	(.53)	(.47)	(.29)		(.14)	(.04)
Primed	556	528	545	—	734	354
	(.74)	(.26)	(.24)		(.01)	(.01)

	Gist task			
			Errors by type	
	Correct	Errors	Don't know	Wrong ans
Normal				
Unprimed	440	503	578	502
	(.79)	(.21)	(.14)	(.07)
Primed	372	420	419	353
	(.94)	(.06)	(.05)	(.01)
Distorted				
Unprimed	412	507	543	608
	(.75)	(.25)	(.16)	(.09)
Primed	378	471	498	459
	(.89)	(.11)	(.05)	(.06)

slower than the undistorted term. Furthermore, the priming does not slow down reading time even for the distorted term. On the contrary, subjects read all words faster when the question was primed. This suggests that in the Literal condition, subjects are especially on their guard to avoid being tricked when the question had been primed. In other words, the distorted term is not read more slowly when primed in the Literal task because it sticks out as inappropriate; subjects are just worried about erring.

GENERAL DISCUSSION

In the introduction, we considered several potential mechanisms to explain the Moses Illusion. The first experiment allowed us to reject a form of the Gricean hypothesis, that giving the undistorted answer to distorted questions was just a case of following conversational postulates. When subjects were instructed to avoid giving the undistorted answer to distorted questions (the Literal task), performance was quite poor. Subjects took much longer to perform the Literal task than the Gist task, and the error rates to distorted questions were almost 50%, suggesting that it is exceptionally difficult to monitor for distortions. As Erickson and Mattson have pointed out, the failure to note the distortions is not typically the result of ignorance: subjects can correctly supply the undistorted term for the question. That is, subjects know that it was Noah, not Moses, who took the animals on the ark.

Another mechanism considered to explain the illusion was *imperfect memory retrievals*. Given that subjects know that they should monitor for distortions, perhaps the information retrieved from memory is incomplete, sometimes omitting the distorted term, thereby explaining the failure to detect the distortion. If the illusion were due to imperfect memory retrievals, then one would expect manipulations that improve access to memory to improve detection. Experiments 2 and 3 made all aspects of the relevant memory trace highly salient. Subjects became very familiar with the critical information and it became highly accessible. When subjects studied the relevant sentence or committed it to memory, answer times dropped significantly and accuracy in terms of answers also dramatically increased. This indicates that the manipulation affected the availability of the correct memory traces. Importantly, the manipulation did not change subjects' sensitivity to the Moses Illusion: Subjects were almost as likely to miss the distorted term when the correct answer had been studied as when it had not been studied; the lower error rate for the studied statements was due primarily to fewer "don't know" responses and fewer wrong answers.

Experiments 4 and 5 were concerned with testing the hypothesis that the locus of the Moses Illusion is in the encoding phase. That is, encoding might be so expectation-driven that people expect to read or hear "Noah" when they begin to process a question that begins "how many animals of each kind . . . ." If the illusion were due to "sloppy" encoding of the target word, then one would expect to see a particular pattern in reading times for distorted target words as a function of whether or not the distortion was detected. Specifically, in the Literal task, subjects should take longer to read the distorted target when the distortion is detected (subject responds "can't say") than when the distortion is not detected and the subject gives the undistorted answer. That is, if expectations drive encoding time and subjects sometimes do not really encode what is there, then they should be fast on the distorted word when they fail to notice that it is distorted.

The data did not support this hypothesis. The relevant comparisons involve reading times for distorted targets in the Literal task, comparing reading times for *correct* and *undistorted answer* times. If anything, subjects read the distorted target faster when they noticed the distortion than when they did not. They read the distorted target 90 ms faster (539 vs. 628 ms) in Experiment 4 when the distortion was noticed than when it was not. In Experiment 5, the same pattern was found for the unprimed questions: subjects read the distorted target 60 ms faster when the distortion was noticed. Even in the situation where subjects studied the relevant statements ahead of time, time spent reading the distorted target did not predict whether the subject would detect the distortion (556 ms if detected, 545 ms if not detected).

Having rejected imperfect retrievals from memory and imperfect encoding of the question as sources of the Moses Illusion, we believe that an imperfect memory match process is the most viable explanation of failure to detect distortions.[5] Conceivably, the entire memory trace is brought back into working memory but not every term is carefully matched to the test question before "reading off" the answer. It seems reasonable that this would be our default process for memory matching. In most situations the form of a question is not likely to closely match the memory representation it queries. Small mismatches would be expected even when the input is a statement rather than a question. So people are accustomed to being tolerant of discrepancies and highly similar terms are allowed to slip by.

If the cause of the Moses Illusion is a partial match process that ignores some dis-

[5] Of course, there may be other explanations of the Moses Illusion not considered here; however, we cannot think of any.

tortions, how would this work? It could not be simply a process where a subset of the words is matched and the rest are completely ignored. If that were true than subjects would false alarm to "How many animals of each kind did Nixon take on the ark?" but they do not (Erickson & Mattson, 1981). On reflection, it seems obvious that the match involves concepts and features rather than words: If specific words were involved in a match, then bilingual speakers would have difficulty answering questions about information that was learned in a language other than the language used for the query. It has recently been shown (Potter & Lombari, 1990) that even immediate "verbatim" retention of a sentence is done at the conceptual level rather than the word level.

For these reasons we believe that the partial match process must be sensitive to the similarity of all the concept words (in the query) to the stored representation of the answer. That is, matching is going on at the feature level, and if the features of the substituted word heavily overlap with the features of the correct target word, then the distortion tends to go unnoticed. Similarity between two terms depends of course on the knowledge state of the person being asked the question. Christopher Columbus does not seem similar to Magellan for Americans, but may seem similar to Belgians.

How did processing change in our experiment compared with "normal" processing? In nonexperimental situations, a person will typically set a loose criterion for finding a relevant memory structure to match an input. In the Literal condition of these experiments, subjects readjusted their criterion upward so as not to be easily tricked. This readjustment was caused both by the explicit instructions to be careful and by subjects' experiences of failure to notice distortions and consequent errors. Even with these reminders, subjects continued to make far more errors as well as to respond more slowly in the Literal condition than in the Gist condition.

What does this *criterion* refer to? We believe that this criterion reflects the *number of tests to accept a match in memory*. When it is very important to be careful, a number of additional tests are probably made between an input probe and a memory structure. For example, all of the relational information in a memory structure might be tested. There is evidence (e.g., Ratcliff & McKoon, 1989) that the relational information is available at a later time than the simple match of features. If the relational information is not tested, the question "On which holiday do children, dressed in costume, go door to door, giving out candy?" would easily slip by as acceptable. Other tests would involve making sure that each word in the query matched the memory structure that contains the answer.

What do these results say about our normal, everyday comprehension of text and other forms of input? These results suggest that our normal mode of processing is as effortless as possible, that people try to do as little work as possible to comprehend or understand. For example, few people would notice the distortion of the expression, "Ask not what you can do for your country, but what your country can do for you." It is less clear whether distortions go unnoticed because it is so difficult to carefully check the input with stored information, or because partial matching is easy and usually works.

Failures to notice distortions occur at all levels: at the phonemic level (e.g., MacWhinney, Pleh, & Bates, 1985; Warren & Warren, 1970), at the syllable or word level, as well as at the phrase level. If failures to detect distortions occur at all levels of comprehension, why should our system operate this way? There are several answers to this. First, the chances are small that the representation of a question will exactly match the representation of the stored information needed to answer it. Second, most distortions are inadvertent and unintentional. It is more sensible to ignore slight mispronunciations than to check for any slight mismatch or violation of expectancy. On the other hand, there are a number of safeguards built into our cognitive system. Large discrepancies from the expected match are detected readily as in the Nixon sample noted above.

Earlier we argued that the conversational postulate seemed an implausible explanation of the Moses Illusion, since subjects found it so difficult to detect distortions even when instructed to do so. However, one could view the tendency to overlook errors, as described above, as simply an operational definition of the conversational postulate where the operation is essentially automatic and not easily influenced by conscious strategies. It is clear that any model that attempts to account for these data must incorporate, at least implicitly, a theory of pragmatics to account for which words can be ignored during the match process. As mentioned earlier, people are sensitive to the *focus* of a sentence and will be much more likely to carefully match and thereby note discrepancies in aspects of the sentence that are in focus, e.g., "It is Moses who took two animals of each kind on the ark" (Bredart & Modolo, 1988). (In contrast, our wh- questions do not focus on the distortion.)

Finally, it is worth noting that people are highly adaptive and become more sensitive to distortions in the sentences once they have been fooled. This is not to say that they do not still find it difficult to monitor for distortions; nevertheless, when apprised of the possibility of distortions and asked to detect them, they make fewer errors: Erickson and Mattson found an approximately 50% illusion rate on questions (computed from those subjects who knew the correct answer). When we compute frequency of illusion by taking the proportion of undistortion divided by one minus the "don't knows" and "wrong answers," we obtain an approximate illusion rate of 36% over all Literal distorted conditions.

APPENDIX: Trivia Questions Used in Experiments

Distorted Target Appears before Undistorted Target; Answers in Parentheses

- How many animals of each kind did *Moses/Noah* take on the ark? (2)
- What country is Margaret Thatcher *president/prime minister* of? (England)
- What kind of tree did *Lincoln/Washington* chop down? (cherry)
- By flying a kite, what did *Edison/Franklin* discover? (electricity)
- What did Goldie-Locks eat at the *Three Little Pigs'/Bears'* house? (porridge)
- What position did Terry Bradshaw play for the Pittsburgh *Pirates/Steelers?* (quarterback)
- Who found the glass slipper left at the ball by *Snow White/Cinderella?* (prince)
- On what holiday to children go door to door, dressed in costume, *giving out/collecting* candy? (Halloween)
- What is the name of the cartoonist who created Mickey Mouse, Donald Duck, and *Pinocchio/Goofy?* (Walt Disney)
- What hero does Clark Kent become when he changes in a *toll/phone* booth? (Superman)
- What is the name of the Mexican dip made with mashed-up *artichokes/avocados?* (guacamole)
- What did *Dasher's/Rudolph's* nose do to help guide Santa's sleigh? (lit up red)
- What is the name of the kimono-clad courtesans who entertain *Chinese/Japanese* men? (Geisha girls)
- What country is famous for cuckoo clocks, chocolate, *stock markets/*

banks and pocket knives? (Switzerland)

- In the biblical story, what was *Joshua/Jonah* swallowed by? (whale)
- In the novel "Moby Dick," what color was the whale Captain *Nemo/Ahab* was after? (white)
- For what scandal was Nixon pardoned unconditionally after Gerald Ford's *election to/assumption of* the presidency? (Watergate)
- The *Wednesday before Easter/first Wednesday of Lent* is called what? (Ash Wednesday)
- From what state was Ronald Reagan a *senator/governor?* (California)
- Who was manager of the World Champion Pittsburgh Pirates in *1980/1979?* (Chuck Tanner)
- What country rescued the hostages from the airport at Entebbe, *Libya/Uganda?* (Israel)
- What is the name of the object whose *circumference/area* is pi-r-squared"? (circle)
- Who began an address with "Four score and *twenty seven* years ago"? (Lincoln)
- What is the name of the carved pumpkin displayed on *Thanksgiving/Halloween?* (jack-o-lantern)
- The Bay of Pigs invasion was orchestrated by *RFK/JFK* against what country? (Cuba)
- At Kitty Hawk, what vehicle was first successfully tested by *Lindbergh/Wright Brothers?* (airplane)
- What is the title of the judge who heads the other *6/8* on the Supreme Court? (Chief Justice)
- What is King Henry VIII of England famous for having *8/6* of? (wives)
- What is the name of the ferocious striped feline found in *Africa/India?* (tiger)
- Who said "Ask not what *you/your country* can do for *your country/you,* but what *your country/you* can do for *you/your country*"? (John F. Kennedy)
- Which month is associated with Mother's Day, *Veterans/Memorial* Day, and spring flowers? (May)
- What phrase followed "To be or not to be" in *MacBeth's/Hamlet's* famous soliloquy? ("That is the question.")
- On what television series did Archie Bunker call his son-in-law *"Meatball"/"Meathead"?* ("All in the Family")
- In what mythology was Venus known as the Goddess of *War/Love?* (Roman)
- What is the name of the well-calibrated stick that was made obsolete by the advent of Texas Instruments' pocket *computer/calculator?* (slide rule)
- When Alexander Haig resigned in protest from President *Ford's/Reagan's* cabinet, what office did he hold? (Secretary of State)
- In which state did General *Grant/Lee* surrender to bring an end to the Civil War? (Virginia)
- In a criminal trial when a man pleads the *Fourth/Fifth* Amendment, who is he refusing to incriminate? (himself)
- What English rock group did the late *Paul McCartney/John Lennon* sing with? (The Beatles)
- What is the name of the once-outlawed *Czech/Polish* labor union? (Solidarity)
- At what farenheit temperature on the *thermostat/thermometer* does water freeze? (32°)
- What baseball team did George Steinbrenner *manage/own?* (New York Yankees)
- What is the name of the man in the red suit and long white beard who gives out *birthday/Christmas* presents from his sleigh? (Santa)
- What is the name of the institution where children stay when their parents did before they are put up for *abortion/adoption?* (orphanage)
- What is the brand name of the *butter/*

margarine that puts a crown on your head? (Imperial)

- How many digits are there in the *zip/area* code required to call another state long distance? (3)
- What winter weather phenomenon is very dangerous to growers of Sunkist *apples/oranges?* (frost)
- What small animal hides the acorns that fall from *elm/oak* trees for his winter food supply? (squirrel)
- What is the name of the gooey substance many children like to spread on bread along with grape *juice/jelly?* (peanut butter)
- By having an apple fall on his head, what did *Galileo/Newton* discover? (gravity)
- What is the name of the famous prize issued by *Denmark/Sweden* for contributions to science? (Nobel Prize)
- Juneau is the capital of what *city/state?* (Alaska)
- When did the *Germans/Japanese* attack Pearl Harbor? (Dec. 7, 1941)
- What year did Thomas Jefferson write the *Constitution/Declaration of Independence?* (1776)
- At what university in Ohio were 4 *National Guardsmen/students* killed during a war protest? (Kent State)
- What statue given to us by *England/France* symbolizes freedom to immigrants arriving in New York Harbor? (Statue of Liberty)
- Gorbachev is the leader of what *capitalist/communist* country? (USSR)
- Snoopy is a *cat/dog* in what famous comic strip? ("Peanuts")
- Who was the first man to fly across the *Pacific/Atlantic?* (Charles Lindbergh)
- Who was the first man to walk on the *sun/moon?* (Neil Armstrong)
- How many *numbers/letters* are there in the alphabet? (twenty-six)
- What bird is known for carrying messages, *gardening/gathering* on rooftops, and defacing statues? (pigeon)

References

Bredart, S., & Modolo, K. (1988). Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica*, 67, 135–144.

Camp, C. J., Lachman, J. L., & Lachman, R. (1980). Evidence for direct-access and inferential retrieval in question-answering. *Journal of Verbal Learning and Verbal Behavior*, **19**, 583–596.

Erickson, T. A., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, **20**, 540–552.

Graesser, A. C., & Murachver, T. (1985). Symbolic procedures of question answering. In A. Graesser & J. Black (Eds.), *The psychology of questions*. Hillsdale, NJ: Erlbaum.

Grice, L. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics:(Vol. 3). Speech acts* (pp. 41–58). New York: Seminar Press (Originally published from William James Lectures, Harvard University.)

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

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processees in reading comprehension. *Journal of Experimental Psychology: General*, **111**(2), 228–238.

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

Lehnert, W. (1977). Human and computational question-answering. *Cognitive Science*, **1**, 47–73.

MacWhinney, B., Pleh, C., & Bates, E. (1985). The Development of sentence interpretation in Hungarian. *Cognitive Psychology*, **17**, 178–209.

Norman, D. A. (1973). Memory, knowledge, and the answering of questions. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington, DC: V. H. Winston & Sons.

Potter, M. C., & Lombari, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, **29**(6), 633–654.

Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology*, **21**(2), 139–155.

Reder, L. M. (1982). Plausibility judgments vs. fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, **89**, 250–280.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, **19**, 90–138.

REDER, L. M., & CLEEREMANS, A. (1990). The role of partial matches in comprehension: The Moses Illusion revisited. In A. Graesser & G. Bower (Eds.), *The psychology of learning and motivation, Vol. 25*. New York: Academic Press.

SCHANK, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 3(4), 552–631.

SINGER, M. (1984). Toward a model of question answering: Yes–No questions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 285–297.

SINGER, M. (1985). Mental operations of question answering. In A. Graesser & J. Black (Eds.), *The psychology of questions*. Hillsdale, NJ: Erbaum.

VAN OOSTENDORP, H., & DE MUL, S. (1990). Moses beats Adam: A semantic relatedness effect on a semantic illusion. *Acta Psychologica*, 74, 35–46.

VAN OOSTENDORP, H., & KOK, I. (1990). Failing to notice errors in sentences. *Language and Cognitive Processes*, 5(2), 105–113.

WARREN, R. M., & WARREN, R. P. (1970). Auditory illusions and confusions. *Scientific American*, **223**, 30–36.

(Received July 30, 1990)
(Revision received December 25, 1990)