Memory for questions and amount of processing

P. N. JOHNSON-LAIRD and C. E. BETHELL-FOX
Centre for Research on Perception and Cognition, Laboratory of Experimental Psychology
University of Sussex, Brighton, Sussex, BN1 9QG, England

Two experiments were carried out on how questions are remembered. Subjects watched a videotape of a series of simple events and then answered 18 questions about these events. The questions were all of the same general syntactic form (e.g., "Did the pencil fall against the jug on A?", where A refers to a particular location). They were designed to elicit three sorts of answer: "yes," "no" because the event took place at another location, and "no" because the event did not take place at all. After the subjects had answered the questions, they were given an unexpected test of their ability to recall them. A difference in the memorability of the questions was predicted on the basis of a procedural theory of comprehension and a hypothesis about memory; subjects should cease to process a question when they realize that it concerns an event that did not take place, and such questions should be harder to remember because they are processed to a lesser degree than the other sorts of question. Experiment 1 confirmed the predictions, but its results in part could be accounted for by assuming that subjects recalled the original events and used them as a cue to remembering the questions. Experiment 2 eliminated this explanation by showing that when subjects do not have to answer certain questions, their recall of them is very poor. However, the same differences in the memorability of the three sorts of question were obtained for both answered and unanswered questions.

The standard version of transformational grammar postulates two sorts of syntactic structure: surface structure and deep structure. Surface structure groups the words and morphemes of a sentence into superficial grammatical constituents such as noun phrases, verb phrases, and so on; deep structure rearranges these constituents so that the fundamental grammatical relations of the sentence, such as its subject and object, are directly identifiable. The semantic representation of the sentence is obtained by combining the meanings of its words according to the deep-structure relations between them. Numerous experiments have been carried out to try to establish that deep structure has a psychological reality and that it is recovered during the process of comprehension (see Fodor, Bever, & Garrett, 1974, for a review). The evidence is equivocal. It can be accounted for in terms of semantics rather than underlying syntax (Fodor et al., 1974, p. 270; Johnson-Laird, 1970, 1974). Indeed, we have proposed that an explicit representation of deep structure is unneccessary for comprehension and speaking (Johnson-Laird, 1977). The cues that would be used to construct deep structure exercise instead an immediate control over the process of semantic representation; comprehension consists in translating a sentence directly into its semantic representation, and likewise, speaking consists in translating a semantic representation directly into a sentence. A

model of this theory has been implemented by Steedman in a computer program that answers simple questions about a one-dimensional universe of discourse in which "particles" move up and down, collide with each other, and so on (Steedman & Johnson-Laird, in press). Each major constituent of a question, roughly speaking, receives an immediate semantic representation in the form of a procedure, and the execution of the procedure leads to a search of memory for the relevant information. For example, the question, "Was Y hit by X at A?" successively sets up the following procedures: EVENT Y OBJ HIT (Find an event consisting of a hitting in which Y was the object); EVENT X SUBJ HIT (and in which X was the subject); EVENT AT A (and which occurred at Location A).

In devising the program, the problem arose as to whether each procedure should be executed as it was set up or, alternatively, stored until the end of the question when all of the procedures could be executed. As far as a computer is concerned, it makes little difference which option is chosen; but what do people do? There is no evidence in the literature, but our intuition was that listeners can start to answer a question as they hear it rather than waiting for its end, and that even when they have decided not to answer, it may be difficult for them to prevent a reply from coming to mind. If someone starts to ask a question, "Does your father . . . ," it is difficult not to bring your father to mind even if you eventually decide that the question is impertinent and refuse to answer it. (There is children's "double bind" that is similarly effective: "Don't think of elephants!") For these reasons, the program was

arranged to execute the procedures as it set them up.

This principle has a consequence that we did not at first appreciate: If a question is based on a false assumption, then once this failure has been revealed, it is pointless to execute any further procedures representing the question, although other procedures may well be run in order to deal with the false assumption. In the case of the question, "Was Y hit by X at A?", then, if in fact Y was *not* hit by X, it is futile to try to determine whether the putative event took place at Location A. Hence, there is a potential saving in executing a procedure as soon as it has been compiled: No further processing of the constituents of a question may be required after the detection of a false assumption. It may not even be necessary to compile any further procedures corresponding to the constituents of the question.

Our theory of comprehension shifts the emphasis from structures to processes. An analogous shift occurred in the study of memory, with the formulation of the "depth-of-processing" hypothesis (Craik & Lockhart, 1972). Current research is much preoccupied with the notion that the deeper the level of processing of an item, the more likely it is to be recalled. For example, responding to the meaning of a word seems to require a deeper level of processing than responding to some aspect of its phonology and leads to a better memory for the word. Unfortunately, as Craik and Tulving (1975) admit, it is difficult to formulate a definition of depth that is independent of the results of memory experiments; and indeed, Nelson (1977) has argued that the hypothesis consequently has no empirical content whatsoever. There appear to be two ways out of the impasse. First, an independent measure of depth of processing might be devised. Unfortunately, the obvious measure, the time taken to process an item, turns out not to have much predictive power (Craik & Tulving, 1975). Second, the hypothesis could be reformulated so as to yield independent predictions. If amount of processing, rather than depth, is defined in terms of the number of decisions that have to be made at a given level in order to carry out the task in hand, then it yields clear predictions about memorability. We have found, for example, that the more semantic components of a word that have to be processed, the better it will be recalled. Subjects had to scan a list of words for those denoting a particular target category such as consumable solids (e.g., "cake"). When the subjects were given an unexpected test of recall, they readily remembered the target words. But they also recalled words with one of the required components, such as "milk" or "coal," better than words with neither component, such as "petrol" (Johnson-Laird, Gibbs, & de Mowbray, 1978).

Although the concept of amount of processing does not yet help to distinguish levels, it does account for those findings that led Craik and Tulving (1975) to

introduce their additional notion of the "richness" or "elaboration" of an encoding. Subjects are more likely to remember the word "shark" after being asked whether it is a type of fish than the word "heaven" after being asked whether it is a type of fish. Craik and Tulving argue that the difference is one of elaboration of encoding, but it can be explained in terms of amount of processing. In order to answer the question about "shark," many aspects of its semantic representation have to be processed (e.g., that sharks swim in the sea, have fins and gills, etc.), whereas in order to answer the question about "heaven," any aspect of its semantic representation suffices to reject it as a type of fish. The theoretical advantage of the amount-of-processing hypothesis is that is it more specific; for example, it predicts that "porpoise" should be better recalled than "heaven," even though both questions would give rise to a negative answer.

Our experiments were designed to test both the procedural model of linguistic performance and the amount-of-processing hypothesis. The former predicts how much processing of a question should occur, and the latter uses this information to predict how memorable the question will be. If the experiments fail, then the reason for their failure will be unclear. If the experiments succeed, however, then the two theories will both be corroborated.

An obvious way to investigate memory for questions is to ask a subject some questions and then to test his or her ability to recall them. But what answers should the subject give to the questions? There are three main sorts of answer that can be given to a yes-no question. Obviously, the same question, "Was Y hit by X at A?" can give rise to all three sorts of answer, but it will be convenient to classify questions in terms of the answers they elicit. Hence, there is what we shall refer to as a "yes" question, to which the answer is "yes." There is a "no" question, to which the answer is of the form: "No, Y was hit by X at C." There is a "false" question, based on a false assumption, to which the answer is of the form: "No, Y wasn't hit by X." There are, of course, other possible answers, but these three exemplify the major categories of interest, since yes and no questions should be better recalled than false questions according to the amount-of-processing hypothesis.

It is desirable to insure that equal numbers of the three sorts of question occur within an experiment, but since the categorization of a question depends, not on its wording, but on its answer, an experimenter needs to know the answers to the questions that he intends to use. Such knowledge might be provided by basing the questions on a story that has already been told to the subjects. An initial pilot study soon convinced us that certain aspects of a story are extremely difficult to control and yet exert a considerable effect on memorability (whether or not as a result of depth or amount of processing is a matter that need not detain us). We

therefore adopted a different technique: Subjects saw
a demonstration in which the experimenter manipulated
a series of objects, such as a pencil, cup, and saucer,
along the lines of Osgood's (1971) procedure. After the
demonstration, the subjects answered questions about it.
Finally, they were given an unexpected test of their
ability to recall the questions they had been asked.
A preliminary study yielded results that reliably
corroborated our predictions, but made use of only a
single demonstration and a single order of presentation
of the questions. Experiment 1 was designed to remedy
these defects.

EXPERIMENT 1

Method

Design. Each subject acted as his or her own control and
was asked 18 questions about a demonstration recorded on
videotape. There were six questions to which the correct answer
was affirmative (yes questions), six questions to which the
correct answer was negative (no questions), and six questions
to which the correct answer was negative because of the false
assumption the questions embodied (false questions). Six
different demonstrations were recorded, and they were
constructed so that the same 18 questions could be asked, but
gave rise to different answers following different demonstrations.
In fact, each of the 18 questions was used equally often in three
categories over all the demonstrations. The subjects were
assigned at random to one of the demonstrations, with the
constraint that each demonstration was used three times in the
experiment as a whole. The questions were presented in a
different random order for each subject.

Materials. Each demonstration involved six items (a pencil,
a pencil sharpener, a cup, a saucer, a jug, and some water that
was initially in the jug) and three locations (a large box labeled
"A," a medium-sized box labeled "B," and a small box labeled
"C"). The experimenter carried out a sequence of actions with
the materials: The pencil was picked up and allowed to fall
against the jug on the large box, the sharpener was picked up
and then brushed against the pencil causing it to fall over, and
so on. The events were devised so as to be of roughly equal
predictability.

The fixed set of 18 questions was related to each demonstra-
tion in a way that yielded six yes questions, six no questions,
and six false questions. Each question contained one verb and
three noun phrases. They were all of the following form: "Did
the pencil fall against the jug on A?" "Did the sharpener push
over the pencil on B?" "Did the cup support the saucer on A?"

The materials were constructed in the following way. Three
blocks of six sentences were made up using 18 different verbs
(9 transitive and 9 intransitive). Within each block, the six
entities in the demonstrations were referred to twice, once by
a subject-noun phrase and once by an object-noun phrase, and
each of the three locations was referred to twice. Each
demonstration was constructed to insure that one block of
questions consisted of yes questions, another block consisted
of no questions, and another block consisted of false questions.
Because there are 3! different ways of establishing this relation,
six different demonstrations were used. The choice of verbs
within a block was aimed to make each of the blocks of
comparable coherence, since each block is of course a set of
yes questions for two of the demonstrations.

Procedure. The subjects were tested individually. The
experimenter told them that they would see a series of events
involving six different entities (which he held up and identified)
in three different locations (which he identified, introducing a

simple mnemonic for them). He also explained that after the
demonstration they would have to answer some questions
about it. The subjects watched one of the videotaped demon-
strations. The experimenter then read out the questions, which
the subjects answered verbally. There was an interval of
approximately 5 sec from the end of one question to the start
of the next one. This provided ample time for the subjects to
respond and for the experimenter to record their answers. After
the subjects had answered the questions, they were taken into
another room and given an unexpected test of their ability to
recall them. They were asked to write down as many of the
questions as possible. They were told that paraphrases were
acceptable and that they could write down their responses in
any order.

Subjects. Eighteen undergraduate students in experimental
psychology at Sussex University took part in the experiment,
which lasted about 25 min. None of them had previously
participated in an experiment of this sort.

Results

The percentages of questions recalled, accurately or
inaccurately, are shown in Table 1; these data are based
solely on the recall of questions that were correctly
answered. There were a further 11 responses that could
not be uniquely related to the original questions. The
predictions were confirmed. Yes and no questions were
reliably better recalled overall than false questions
(Wilcoxon's T = 5.0, N = 17, p < .005, one-tailed). The
difference between the overall recall of yes and no
questions was not reliable (Wilcoxon's T = 35, N = 16,
p > .05, two-tailed). However, it is clear that yes
questions tended, if recalled, to be recalled accurately
(Wilcoxon's T = 7, N = 14, p < .01, two-tailed), whereas
no questions showed no such bias and were often recalled
inaccurately (Wilcoxon's T = 14, N = 10, p > .05,
two-tailed). This interaction was reliable (Wilcoxon's
T = 20, N = 16, p = .01, two-tailed). The inaccurate
recalls consisted largely (79%) of remembering the
body of the sentence correctly but the locative phrase
incorrectly. In the case of no questions, this location
corresponded with the actual location of the event
instead of that referred to in the question in 24 out of
the 30 such errors.

There was a rather large proportion of errors made
in answering the questions (19.7%). Errors occurred
reliably more often with false questions (34.3%) than
with the other two sorts of question (12.9% for yes
questions and 12.0% for no questions); 12 subjects
conformed to this pattern and only 4 went against it
(sign test, p < .05). When the recall of such questions

Table 1
Percentages of the Three Sorts of Question Recalled Accurately
(Verbatim or in Paraphrase) or Inaccurately in Experiment 1

	Kind of Question		
	Yes	No	False
Accurately Recalled	41.5	16.8	11.3
Inaccurately Recalled	26.6	37.9	15.5
Overall	68.1	54.7	26.8

is included in the main results, their overall pattern remains the same. However, it is noteworthy that 45.4% of false questions incorrectly answered "yes" were recalled, whereas only 26.8% of false questions correctly answered "no" were recalled (Wilcoxon's T = 7.5, N = 10, p < .05, two-tailed).

Discussion

The results corroborate our hypothesis that the greater the amount of processing of a question, the more likely it is to be remembered. A single video presentation of a demonstration evidently did not lead to a very good performance in answering the questions, and the inaccuracy of the answers to some extent weakens the findings. Moreover, subjects were prone to recall the no questions inaccurately and to substitute the actual location of the event for the one referred to in the question. This finding suggests an obvious alternative explanation of the results. Subjects remember the demonstration and use it to guide their attempts to reconstruct the questions. Such a strategy would naturally account for both the failure to recall false questions (they have no counterpart in the demonstration) and the erroneous recall of no questions. It was therefore necessary to eliminate this alternative explanation if we were to substantiate our hypothesis.

EXPERIMENT 2

Method

Design. In this experiment there was an additional independent variable of whether or not a question had to be answered. It was introduced in order to discover whether memory for questions is determined by amount of processing or by memory for the events they concern. The design of the experiment was identical to that of the previous experiment, except that the subjects had to answer only 12 of the 18 presented questions. They were told that they must not answer the remaining six questions. The instruction was given after the demonstration and it was done by naming one of the six objects featured in the demonstration: No question referring to this object was to be answered. This procedure insured that for each subject there were two yes questions, two no questions, and two false questions that were not to be answered. The relevant object was selected at random with the constraint that each object be used three times in the experiment as a whole but no more than once for any particular demonstration.

We made two predictions based on the amount-of-processing hypothesis. First, questions that were answered would be better recalled than questions that were not answered. This prediction cannot be derived from the assumption that subjects are merely recalling the demonstrations. Second, since it is impossible to

repress all processing, there would be the same pattern as before in the recall of both answered and unanswered questions: Yes and no questions would be better recalled than false questions.

Materials. The demonstrations of the previous experiment were slightly modified in order to render them more coherent; that is, the sequences of actions moved less haphazardly from one location to another. The same six entities and three locations were featured, and again one set of 18 questions was devised to apply to all the demonstrations. The majority of questions used the same verbs as in the previous experiment, but three new verbs were introduced. There were nine transitive verbs and nine intransitive verbs. The relations of the blocks of questions to the different demonstrations were exactly the same as in Experiment 1, insuring that each question was used as a yes question, a no question, and a false question for two different demonstrations.

Procedure. The subjects were tested individually. The same basic procedure was used as before, except that the subjects viewed the videotape twice. Likewise, the experimenter explained that certain questions referring to a "mystery" object were not to be answered; the subject should simply respond, "I can't answer this." After the second presentation of the videotape, the experimenter identified the "mystery" object and then read out the questions in the usual way. On the few occasions that subjects did answer a question about the "mystery" object, they were reminded that they had been instructed not to answer such questions. The whole session was recorded on tape in order to allow the latencies of the subjects' replies to be measured.

After the presentation of the questions, the subjects were taken into another room and given an unexpected test of their ability to recall all of the questions, including those that they had not been required to answer. They wrote down their answers in the same way as in the previous experiment.

Subjects. Eighteen undergraduate students from the same population as before were paid to take part in the experiment. None of them had previously participated in an experiment of this sort.

Results

The percentages of answered and unanswered questions recalled, accurately or inaccurately, are shown in Table 2. These data are based only on questions that were responded to correctly; there were only a small proportion of errors in answering questions (8%). The results confirmed the first prediction: Questions that are answered are better remembered than questions that are not answered. This difference was reliable both overall (Wilcoxon's T = 12, N = 17, p < .005, one-tailed) and for accurate recalls (Wilcoxon's T = 18.5, N = 17, p < .005, one-tailed), but it was not reliable for inaccurate recalls (Wilcoxon's T = 39.5, N = 15, p > .05, one-tailed).

The recall of questions that were answered confirmed the second prediction: Yes and no questions were better

Table 2
Percentages of Questions Recalled Accurately or Inaccurately in Experiment 2

	Questions that Subjects Had to Answer			Questions that Subjects Were Not to Answer		
	Yes	No	False	Yes	No	False
Accurately Recalled	47.1	11.6	7.2	26.7	3.0	3.4
Inaccurately Recalled	33.8	42.0	14.5	23.3	33.3	10.3
Overall	80.9	53.6	21.7	50.0	36.3	13.7

recalled than false questions (Wilcoxon's T = 5, N = 17, p < .005, one-tailed). Also, yes questions were recalled accurately, whereas no questions showed no such bias and were often recalled inaccurately, an interaction that was significant (Wilcoxon's T = 9.0, N = 14, p < .005, one-tailed). The recall of questions that were not answered yielded a comparable pattern of results: Yes and no questions were better recalled than false questions (Wilcoxon's T = 8.0, N = 14, p < .005, one-tailed), and there was the same interaction between accuracy of recall and sort of question, with the yes questions tending to be recalled accurately and the no questions showing no such bias (Wilcoxon's T = 19.0, N = 13, p < .05, one-tailed). There was no effect on the recall of unanswered questions of whether the "forbidden" item occurred as the subject or object of the sentence. In the case of accurate recalls, the forbidden item was the subject in five cases and the object in three cases. Twenty of the recalled sentences could not be uniquely identified with one of the original questions (12 were answered questions and 8 were unanswered questions).

The tape recordings were analyzed using a speech spectrometer (Jeffrey & Longuet-Higgins, Note 1), and the latencies of the subjects' answers to the questions were measured with an accuracy of ±10 msec from the start of the locative noun ("A," "B," or "C") of the question. The mean latencies to respond were as follows: 1.15 sec for yes questions, 1.41 sec for no questions, and 1.17 sec for false questions. There was no reliable difference between these latencies [Friedman $\chi_r^2(2) = 2.1$, p > .03], and the apparent tendency for no questions to yield larger latencies was not significant (Wilcoxon's T = 46, N = 18, p > .05). There was, however, a significant effect of practice on the latency of answering the questions: The mean latency for the first block of six trials was 1.42 sec, for the second block of six trials, 1.22 sec, and for the last block of six trials, 1.18 sec (Page's L = 235, p < .001).

GENERAL DISCUSSION

The results of the experiments corroborate both the procedural model of linguistic performance and the amount-of-processing account of memorability. Once listeners have discovered that a question is based on a false assumption, they do not execute any further procedures corresponding to the meaning of the question; their memory for such a question is correspondingly impaired, since they have processed it to a lesser degree than an ordinary yes question. The recall of the no questions, however, suggests that our initial formulation of the amount-of-processing hypothesis was oversimplified. When subjects are asked a question of the form, "Did X hit Y at Location A?" and reply correctly, "No" (because X hit Y at Location B), their recall of the question is likely to suffer in two ways.

First, they are not likely to remember the question as well as one to which the answer was "yes." Second, if they recall the question at all, there is a fair probability that they will recall it inaccurately, as "Did X hit Y at Location B?" Even in the case where the subjects did not have to reply to a no question in Experiment 2, six out of the nine locational errors that were made were of this form, and there were only two other inaccurate recalls of such "no" questions. It is therefore plausible to assume that, although the amount of processing of a question is an important determinant of how likely it is to be recalled, the content that is recalled is likely to depend on the outcome of this processing.

The simplest way to explain the revision in our thinking is to consider again the operation of the computer program (Steedman & Johnson-Laird, 1978). The reader will recall that it represents a question such as "Did X hit Y at A?" as a series of search procedures that scan memory for such an event. If each search is satisfied, a corresponding set of assertions is returned, such as: [EVENT 21 X SUBJ HIT] (Event 21 consists of a hitting in which X was the subject); [EVENT 21 Y OBJ HIT] (and in which Y was the object); [EVENT 21 AT A] (and which occurred at Location A). However, if the event occurred at some other location, say, B, then the last of the searches will fail, and the program will then call a simple procedure to determine the event's correct location. The resulting set of assertions, such as [EVENT 21 X SUBJ HIT], [EVENT 21 Y OBJ HIT], [EVENT 21 AT B], is used to construct a "helpful" answer: "No, X hit Y at B." If the first search procedure fails because X did not hit anything, then the subsequent searches will not be carried out, and the program will reply: "No, X did not hit Y." Our results suggest that people perform in an analogous fashion and, in particular, that they are likely to remember those "assertions" returned as a result of searching their memories. It follows, of course, that they were likely to recall the no questions with the actual location of the event substituted for the questioned location, and that they are likely not to recall a false question at all.

Could these phenomena be explained by some entirely different theory? One possibility is that subjects use their memory for the actual events as a cue for recall, but this alternative is ruled out by the poor recall of unanswered questions; if subjects were simply remembering the events, then whether or not a question is answered should be immaterial to its recall. The process of answering a question must at the very least provide some information that strengthens memory for the events or that can be used to select actual questions from the candidates generated by recalling the events. Otherwise, subjects would recall answered and unanswered questions equally well and would even "recall" questions that were not actually asked; that is, they would confabulate by constructing questions corresponding to states of affairs in the demonstrations

for which no questions were framed. There were at least 10 such questions that each subject could have constructed in Experiment 2, but only two responses were observed that could have been constructed in this way. The strongest alternative hypothesis is, accordingly, that the process of answering a question strengthens only those events that correspond to its constituents, and that as soon as subjects hear the name of an object signaling that a question is not to be answered, there is some probability that they stop processing the question. Hence, answered questions will be better remembered than unanswered questions, and both varieties will give rise to the same pattern of results for yes, no, and false questions. There is one crucial datum that will decide between this account and our proposals: We need to examine memory for false questions that are incorrectly answered.

If a false question elicits the incorrect answer, "yes," then, according to our hypothesis, the subject has been able to infer from the contents of his memory that each of its constituents is satisfied. (Of course, he may simply have guessed the answer, but we shall ignore this possibility on the grounds that it has no material effect on the following prediction.) Such a process clearly takes a greater amount of processing than that which occurs in generating the correct answer, "no," to a false question. It follows that incorrectly answered false questions should be better remembered than correctly answered false questions. However, this prediction cannot be made by the hypothesis that questions strengthen memory for the actual events because, by definition, there are no such events for a false question. There were sufficient errors in Experiment 1 for the comparison to be made, and, as we reported, there was a reliable tendency for the incorrectly answered false questions to be recalled more often than the correctly answered false questions. Experiment 2 yielded the same pattern of results, but the figures were too small for statistical comparison (two out of the three incorrectly answered false questions were recalled).

Our theory is incompatible with the proposal that syntactic and semantic processes are autonomous. According to this view, the effects of manipulating syntactic and semantic complexity are additive (Forster & Olbrei, 1973). In fact, our computer program for answering questions operates in flagrant breach of this principle, and we have reported elsewhere some preliminary experimental results that suggest that human beings do so, too (Steedman & Johnson-Laird, 1978). Autonomy is violated because the program builds up semantic procedures, and executes them, as it progresses through a sentence. It uses semantic information to help it with syntactic decisions. Our findings can be interpreted to suggest that subjects do, indeed, control their processing of one part of a question in terms of the outcome of the processing of its earlier parts, and this result is plainly compatible with the interactive nature of our general theory of linguistic performance (Johnson-Laird, 1977).

In conclusion, the present findings suggest that the likelihood of recalling a question depends on how much it is processed and that the content of what is recalled is determined by the outcome of the process that compares the question with the memory for the events, whether that memory is accurate or inaccurate. What do you remember of a question? The short answer is: its answer.

REFERENCES

CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* 1972, 11, 671-684.

CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General,* 1975, 104, 268-294.

FODOR, J. A., BEVER, T. G., & GARRETT, M. F. *The psychology of language.* New York: McGraw-Hill, 1974.

FORSTER, K. I., & OLBREI, I. Semantic heuristics and syntactic analysis. *Cognition,* 1973, 2, 319-347.

JOHNSON-LAIRD, P. N. The perception and memory of sentences. In J. Lyons (Ed.), *New horizons in linguistics.* Harmondsworth, England: Penguin, 1970.

JOHNSON-LAIRD, P. N. Experimental psycholinguistics. *Annual Review of Psychology,* 1974, 25, 135-160.

JOHNSON-LAIRD, P. N. Psycholinguistics without linguistics. In N. S. Sutherland (Ed.), *Tutorial essays in psychology* (Vol. 1). Hillsdale, N.J: Erlbaum, 1977.

JOHNSON-LAIRD, P. N., GIBBS, G., & DE MOWBRAY, J. Amount of processing and memory for words. *Memory & Cognition,* 1978, 6, 372-375.

NELSON, T. O. Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior,* 1977, 16, 151-171.

OSGOOD, C. E. Where do sentences come from? In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology.* Cambridge: Cambridge University Press, 1971.

STEEDMAN, M. J., & JOHNSON-LAIRD, P. N. A programmatic theory of linguistic performance. In P. T. Smith & R. N. Campbell (Eds.), *Proceedings of the 1976 Stirling conference on the psychology of language.* London: Plenum, in press.