

1976

## Review effects of adjunct questions on learning from prose/

Susan Bennett Sefkow  
*University of Massachusetts Amherst*

Follow this and additional works at: <https://scholarworks.umass.edu/theses>

---

Sefkow, Susan Bennett, "Review effects of adjunct questions on learning from prose/" (1976). *Masters Theses 1911 - February 2014*. 1948.  
<https://doi.org/10.7275/6871555>

This thesis is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses 1911 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).



UMASS/AMHERST



312066013692293



REVIEW EFFECTS OF ADJUNCT QUESTIONS  
ON LEARNING FROM PROSE

A Thesis Presented

By

Susan Bennett Sefkow

Submitted to the Graduate School of the  
University of Massachusetts in partial  
fulfillment of the requirements for the degree of

MASTER OF SCIENCE

February

1976

Psychology


REVIEW EFFECTS OF ADJUNCT QUESTIONS  
ON LEARNING FROM PROSE

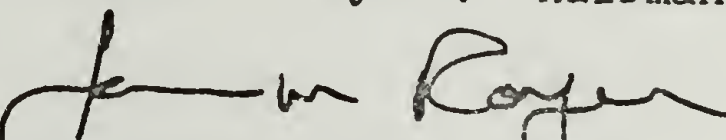
A Thesis


By

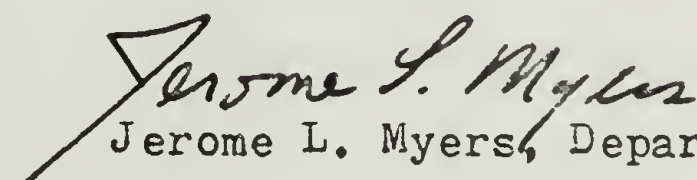
Susan Bennett Sefkow

Approved as to style and content by:

  
Jerome L. Myers, Chairman of Committee

  
James M. Royer, Member

  
Miriam G. Rogers, Member

  
Jerome L. Myers, Department Head  
Psychology Department

February 1976

## Acknowledgements

I wish to thank the members of my thesis committee, including Chuck Clifton whose membership was interrupted by his sabbatical, for their thoughtful comments and criticisms.

I particularly want to recognize Jerry Myers who was an invaluable source of insightful suggestions, encouragement, and patience.

Finally, I wish to express special thanks to my husband, Tom, for his patient assistance and encouragement throughout this investigation.

This research was supported in part by Grant MH-25434 from the National Institute for Mental Health to Dr. J.L. Myers. Funds for data analysis were provided by a grant from the University of Massachusetts Computing Center.

## ABSTRACT

REVIEW EFFECTS OF ADJUNCT QUESTIONS  
ON LEARNING FROM PROSE

(February 1976)

Susan Bennett Sefkow, B.A., Yale University

Directed by: Professor Jerome L. Myers

Two experiments were performed to (a) determine whether the processing initiated by an attempt to answer a question about information available only in memory could substantially facilitate retention of that material, (b) identify the locus of such a review effect, and (c) begin to examine the nature of the process responsible for the effect. Ss listened to five prose passages and immediately after each were asked to verify either a true inference drawn from the passage or a false statement. Subsequent free recall data, collected under both incidental and intentional learning instructions, demonstrated the existence of a review effect (true-probed passage recall exceeded false-probe recall) and indicated the strength of the effect to be centered on those relations comprising the inferences (a mean advantage of up to 24.8% was obtained over relations from both true- and false-probed passages which were irrelevant to the inferences). When Ss listened to the passages and were then given the inferences exclusively as retrieval cues at the time of recall, the effect disappeared. This suggests that the backward review effect

can not be attributed solely, or even substantially, to a cueing or retrieval phenomenon but rather to a strengthening or integration of the memory traces at the time of the probe.

## Table of Contents

Acknowledgements .....	iii
Abstract .....	iv
List of Tables .....	vii
List of Figures .....	viii
Introduction .....	1
History of the Problem .....	4
Backward Review Effects: A Heuristic Framework ..	13
Generale Procedure and Rationale .....	17
Experiment I .....	22
Method .....	27
Results .....	31
Discussion .....	47
Experiment II .....	52
Method .....	55
Results .....	56
Discussion .....	68
General Discussion .....	71
References .....	75
Appendices .....	79
Appendix A: Passages .....	79
Appendix B: Instructions .....	82



## List of Tables

Table		Page
1	Experiment I: Five Inferential Classes for Each of Five Passages .....	28
2	Experiment I: Mean Percentage of Links and Inferences Recalled as a Function of Passage Serial Position, Averaged Across All <u>S</u> s and Passages .....	34
3	Experiment I: Mean Percentage of Inferences Recalled and Implied Per <u>S</u> .....	36
4	Experiment I: Mean Percentage of Total Links and Inferences Recalled Per <u>S</u> .....	38
5	Experiment I: Mean Percentage (Adjusted) of Links Recalled Per <u>S</u> .....	41
6	Experiment I: Mean Percentage of Inferences Recalled Per <u>S</u> .....	43
7	Experiment II: Mean Percentage of Inferences Recalled and Implied Per <u>S</u> .....	59
8	Experiment II: Mean Percentage of Total Links and Inferences Recalled Per <u>S</u> .....	60
9	Experiment II: Mean Percentage (Adjusted) of Links Recalled Per <u>S</u> .....	64
10	Experiment II: Mean Percentage of Inferences Recalled Per <u>S</u> .....	66

## List of Figures

Figure		Page
1	Experiment I: Mean True-False and Stated- Implied Recognition Ratings as a Function of Item Type .....	48
2	Experiment II: Percentage of Links and Infer- ences Recalled per Passage (minus cues) as a Function of Probe Type .....	62

## Introduction

The question of how humans learn from spoken or written discourse has pervaded the field of experimental psychology for many decades. The usual approach of verbal learning theorists has been to examine learning tasks using simple materials; e.g. paired associate, serial, or verbal discrimination learning employing nonsense syllables, numerals, or single words. Their goal has been to control learning situations to the extent that the salient characteristics of this complex process can be isolated from the myriad of possible variables and, in turn, their interactions examined by a systematic recombination of these variables. In this way it is hoped that a comprehensive theory of verbal learning can be constructed from a firm foundation of basic research.

Another school of thought, gaining momentum during the last few decades, seriously questions how well this "distillation process" will explain the learning processes occurring in more realistic settings. The basic question is: Will the concepts developed using artificial materials necessarily generalize to prose? After all, the great majority of information that most humans encounter and try to assimilate takes the form of discourse, written or oral. One need only consider that prose has a much more compelling sequential nature and complex organization than



a list to realize that further research is necessary to relate "list learning" phenomenon to prose.

This problem has become the concern of both basic researchers interested in constructing comprehensive theories of learning and memory and those in education seeking practical applications of verbal learning research. There is no doubt that there exists a need to bring about a better relationship between traditional verbal learning approaches (most often identified with "basic" research) and "real life" content and context (too often only the concern of "applied" areas). To pursue this does not necessarily imply the abandonment of experimental control as frequently suggested; it just makes the task more difficult. The research presented here was designed to contribute to just such a liaison.

Rothkopf (1972), whose interests are in the applied area, is one proponent of switching the emphasis in research on human learning from lists to prose. Furthermore, he questions whether variables found to be potent in more traditional paradigms will be "relevant to the really big effects in the prose domain." Rothkopf's objections focus primarily on the unrealistic constraints that have been placed on Ss. Under normal circumstances, a person asserts considerable control over the frequency and rate of inspection of information he wishes to learn

(particularly when in written form). This is certainly not the case in the usual laboratory setting. Moreover, the quantity and complexity of verbal material typically dealt with far exceeds usual experimental requirements. From this perspective, Rothkopf has proposed that control processes should be at least as important as structural variables when predicting learning from prose. He has gone on to generate an area of research which, in fact, has demonstrated that learning outcomes can be altered by experimental manipulation of Ss' control processes, though most often the effects have been quite small. This research will be reviewed shortly, but in general this manipulation has been done indirectly by embedding statements or questions in various positions within textual materials. It has been suggested that these inserted items govern the Ss' inspection of subsequent materials, thereby affecting its later recall. This explanation relies heavily upon the principles of operant conditioning and in doing so tends to focus on events external to the Ss. It is felt that this approach is of limited use in explicating the nature of such an effect. The research reported here has adopted a more cognitively oriented framework, refocusing on the Ss' internal processes. Its intent then was to begin to develop a paradigm, with respect to this viewpoint, within which the facilitative effect could be

more carefully examined. While acknowledging the likelihood of forward operations, it was specifically directed towards investigating another factor which may contribute substantially to such an effect. This auxiliary or possibly alternative approach centers around whether or not inserted questions could actually be used to enhance memory for materials already encountered but no longer directly available for further inspection. In other words, the major question of interest was: Does the processing initiated by an attempt to answer a question about information available only in memory facilitate retention of that material?

### History of the Problem

Recent research has focused on effects on learning, both incidental and intentional, of questions interspersed within prose materials (Rothkopf, 1966; Frase, 1968a; Rothkopf & Bisbicos, 1967; etc.). It has been consistently shown that questions inserted before the text segment from which they were derived result in increased retention of content relevant to those questions. Furthermore, questions inserted after such a text segment result not only in better retention of content specific to those questions, but also better retention of nonspecific or incidental material. This holds true even when Ss are not



allowed to reinspect passage segments nor are given feedback as to the correct answers to the adjunct questions.

Investigators have attempted to explain both this pre- and post-question induced facilitation in terms of forward operations: that is, the questions primarily affect processing of the succeeding, not preceding text. This will be clarified as the following discussion presents the various arguments focusing on forward moving effects.

Rothkopf (1966) has attributed the effects of adjunct questions to two sources: first, the adjunct questions are responsible for direct review; i.e., specific instructive effects, in that performance is enhanced on the identical questions when they are encountered as criterion items. Second, the inserted items serve to control the reader's "mathemagenic" behaviors, defined as any activity that results in the reader attaining specific instructional objectives. These mathemagenic or inspection behaviors can take such forms as attention, learning to learn, and set. It is this second concept that Rothkopf asserts is particularly responsible for the increased retention of incidental material in the post-question condition. The reader starts out with some set of inspection behaviors and then encounters an inserted post-question. If he successfully answers it, this set of behaviors is reinforced and thus

maintained through the next segment. However, if the question is not correctly answered, the reader modifies the way in which he processes the following text. In this manner, the set of adjunct questions serves to gradually shape or refine the reader's problem solving activities. Clearly, on the basis of this hypothesis it can be predicted that, relative to a control group, the probability of correctly answering either an adjunct post-question or a criterion test item should be a nondecreasing function of the item's ordinal position within the text. In fact, we shall see that in the few experiments (e.g., Frase, 1968c; Watts & Anderson, 1971) where such functions were reported their shapes did not satisfy this requirement. These results begin to cast doubt on Rothkopf's forward shaping hypothesis as it is formulated and begins to point to the need for a more satisfactory explanation(s). Subsequently, this issue will be dealt with more fully.

Frase (1968b) has supported Rothkopf's contentions that both pre- and post-questions, dealing with criterion relevant material, facilitate retention of this material through transfer of training and also that post-questions assert a forward operating control over inspection behaviors. Additionally, Frase proposed that pre-questions actually put limits on what a reader processes; that, in fact, retention of incidental information is actually

depressed as the reader "focuses" only on those stimuli relevant to answering the pre-questions.

A study by Rothkopf and Bisbicos (1967) was one of the first to manipulate question type and found that, indeed, it affected the kind of information whose recall was facilitated. For example, one group of Ss received adjunct post-questions dealing with quantitative terms and, in turn, scored significantly higher on the criterion test on that kind of information (in relation to the other groups). The effect appeared to be stronger when the information came towards the end of the text. These results were taken to imply that certain materials could be selectively facilitated as a function of adjunct question type, as well as support for a forward shaping hypothesis.

Typically, the questions used in this research have dealt with such specifics as names, dates, and quantities lifted directly from the passages and thus have called for little more than rote recall. In fact, the term "recall" implies more than was often required of Ss: "rote recognition" would be more apropos in many cases. Watts and Anderson (1971) questioned the practical value of this research for just that reason. After all, verbatim recall is of little use if the concept it represents is not comprehended or if it can not be applied in novel circumstances. Few would argue against the latter being more



relevant to educational goals.

Watts and Anderson decided to investigate whether or not post-questions could be used to promote comprehension of written discourse. The goal was to try to force "deeper processing" by use of adjunct questions and therefore bring about more potent, as well as meaningful, retention of that material. Their study contrasted post-question conditions calling for a psychologist's name, a repeated example of a psychological principle, a new example of that principle, and a no-question condition. The group which had to apply the principles to new examples performed significantly better on the criterion test than all other groups. Additionally, those Ss answering name questions recalled even less than the no-question control Ss. The data clearly supported the notion that question type is an important variable influencing the nature of the facilitation.

Another analysis of the retention scores as a function of the ordinal position of passages across conditions showed a classical serial position effect. This was contrary to the forward shaping hypothesis which predicts increasing retention of incidental material for succeeding units as the reader's processing activities become better adapted in the post-question condition. Frase (1968c) also failed to find this particular question type (pre versus

post) by paragraph position interaction. Thus, he concluded that post-questions do operate in a "forward" manner, but instead of changing inspection behaviors, they simply confirm and maintain initial skills. In the same study, Frase found that post-questions were superior to pre-questions even on the first paragraph. He attributed this to the focusing or suppressive effect of pre-questions. As an alternative explanation of these data, Watts and Anderson (1971) suggested that perhaps the facilitation was the result of some backward review activity. That is to say, in seeking an answer to a post-question, a memory scan of the relevant passage may be initiated in order to locate information necessary to answer the question. Depending upon what has been stored in memory and how, as well as the nature of the scan and the requirements of the question, incidental material would also somehow be reviewed, rehearsed, or further processed, resulting in its enhanced retention.

A recent study by McGaw and Grotelueschen (1972) focused on ascertaining the existence of a backward review effect. They argued that if some kind of backward review was operating then one would expect greater facilitation for criterion test items whose content was closely related to the inserted post-questions than for those criterion items which were not. Though no basis for this prediction

was presented, it implies some sort of selective or partial scan of memory. Another prediction was that, presumably because of memory decay, this facilitation would be stronger for short preceding lags and weaker for longer lags. McGaw and Grotelueschen investigated these possibilities by manipulating textual distance of the inserted questions and their semantic relationship to criterion test items. In confirming their predictions, evidence was presented in favor of a backward review, along with increased attentiveness and forward shaping of inspection behaviors following the adjunct questions. The nature of the backward effect, whether it was a retrieval phenomenon or the result of a change in storage, was not evident. The investigators suggested two variables that might be of importance in explaining a backward review: similarity of semantic content of the inserted to the criterion test questions and physical proximity of the inserted questions to the relevant text. They also suggested that the inserted questions may just be serving as response cues for the criterion questions. This hypothesis will be specified more carefully and investigated in the present research.

Rothkopf and Billington (1974) designed a study to replicate the results of McGaw and Grotelueschen. Since the findings of these studies are of import, they merit closer attention. Because the same materials were used and



the paradigms for both studies were similar, a description of the Rothkopf and Billington study will be presented. A 24 page (approximately 6000 words) prose passage was used. Two factual, completion-type questions were constructed from each page such that the items dealt with the same content, yet could not be answered as a function of inspection of the other. Typically, the questions were verbatim sections of the passage with appropriate portions deleted to form the questions. Hence, the matched pairs had many words and phrases in common.

The questions were divided into three groups: inserted post-questions, an immediate criterion test (CT1), and a second criterion test (CT2). CT1 was composed of items matched to each of the inserted questions along with an equal number of new items. CT2 was formed by repeating the inserted questions along with items matched to the new CT1 questions.

Ss were run in small groups and were allowed to control their study and test times, as done in earlier studies. They did not receive feedback for the adjunct questions.

The results confirmed the previous findings of McGaw and Grotelueschen (1972): the average number of correct responses on matched items ( $\bar{X}=3.58$ , out of a possible 12) in CT1 was greater than for unmatched items ( $\bar{X}=3.17$ ) across

conditions. For CT2, the average number of correct responses for previously seen items ( $\bar{X}=5.70$ ) was greater than for unseen items ( $\bar{X}=3.24$ ). The differences, though quite small, were significant at the .01 level.

The fact that the adjunct questions facilitated performance on criterion test items matched for closely related content does support the existence of a backward effect. The Rothkopf and Billington data additionally suggest that the time interval between encoding of the textual material and the subsequent encounter with the matched inserted question may be of importance to the facilitation process. First, facilitation was greatest when the inserted question tapped content from the immediately preceding page; the effect weakened as this distance increased. Second, there were no signs of improved performance on items in CT2 matched to the new items in CT1. In other words, questions first encountered during CT1 did not elevate performance on matched questions in CT2. It would seem that the state of memory for the passage is crucial to the degree of influence which inserted questions might have over the passage's retention. For information to be enhanced in memory, or for additional schemes for retrieval of it to be acquired, the information must be accessible. If it has already been lost or faded in memory, there is nothing to enhance.

Clearly, evidence now exists which suggests that retention of prose can be enhanced by probing memory with questions about the material. In both of the above studies, this effect was significant but small (about a 3.5% improvement). In general, small effects are the rule for research involving inserted questions. Ladas (1973) has recently shown that, in fact, several studies exhibiting small differences between conditions (e.g., Rothkopf, 1966; Rothkopf and Bisbicos, 1967) rest their arguments on improper statistical analyses. Type I error rates were inflated, resulting in reported significance where there actually was none. By adopting the working hypothesis that a backward review effect is viable, the present research focused on ascertaining whether or not more sizeable effects could be achieved.

#### Backward Review Effects: A Heuristic Framework

Recent findings cited in the verbal learning literature speak from a processing viewpoint to the conceptions formulated in attempts to explain the effects of inserted questions. It has been demonstrated repeatedly (e.g., Hyde & Jenkins, 1969; Johnston & Jenkins, 1971; Till & Jenkins, 1973) that within various list learning paradigms incidental and intentional memory for items depend upon the level or type of processing given to an item at presen-



tation. For example, having a S categorize a word semantically produces better delayed recall than having him perform a more superficial analysis; e.g., noting some physical attribute of the item. Therefore, the degree of processing and hence later recall varies in response to the task demands placed on Ss' processing capabilities. The nature of the materials being dealt with is also of issue: the recall advantage enjoyed by semantically processed words was magnified when the lists were of highly associated words (Hyde & Jenkins, 1973). The applicability of these results to prose research is promising. Mistler-Lachman (1974) has already extended them to sentences.

Identical pre- and post-questions differ in their effect on memory for prose. Why? Within a "processing" framework, it is not difficult to recognize some diverging characteristics of their processing requirements. Let us consider the more immediate requisites. On the one hand, Ss know they can do little more than guess at the answers to pre-questions. They have not as yet had access to the information being tapped and a correct answer, at least for the moment, is not required nor is it to be expected. On the other hand, Ss encountering post-questions are exposed to the relevant information first. This appears to be an important distinction. Ss experiencing post-questions face a very different task than their counterparts. An actual

response is required of them. To fulfill such expectations the Ss must rely on what they have stored in memory. Herein lies the possibility of a meaningful backward review effect, not feasible in pre-question conditions. The present research was designed to investigate just such a possibility.

While the position of an inserted question is apparently critical in determining its processing requirements, other features are of at least equal importance. Investigators (Rothkopf & Bisbicos, 1967; Watts & Anderson, 1971; Rickards & Di Vesta, 1974) have already documented the fact that questions asking for verbatim recall/recognition of specific items are less effective "enhancers" than those involving more meaningful, integrative learning. This point raises two distinct, though not necessarily independent, issues that need to be explicated. One concerns the type of question. Does it require rote memorization of facts or what is termed comprehension? The latter logically entails more thorough processing than the former, and thus, in line with the orienting task findings, predicts the confirmed result: better retention. The mode or form of an adjunct post-question; that is whether it requires recognition or recall, is also important. It is generally acknowledged that recognition is somehow easier than recall. Successful recognition is generally thought

to depend upon the target item's availability (strength) while recall, in addition, relies on its accessibility (cf. Kintsch, 1970). Therefore, it can be predicted that post-questions involving recall, as opposed to recognition, will engender more complex processing and hence better retention.

The nature of the criterion test is also important to consider. A point just raised is relevant: recognition and recall measures appear to tap different aspects of memory. Furthermore, research exists which suggests that recall and recognition measures vary in their ability to detect the effects of a variety of independent variables. It has been proposed that manipulations involving organizational variables have their effects on recall and little if any on recognition, while just the opposite appears to be the case with those involving exposure duration (Kintsch, 1970; McCormack, 1972; Woodward, Bjork, and Jongeward, 1973). If the processing initiated by inserted post-questions involves integration of information, elaboration, etc., then a criterion test requiring recall may best reflect the effects of such a process. On the other hand, a recognition test may be called for if the backward review entails some kind of simple scan or review. The implication is that attention should be paid to the form of the criterion test, not only because of differences in overall sensitivity to



memorial representations, but because this sensitivity may well extend to the detection of effects due to the inserted questions.

The purpose of the above discussion was to provide the reader with some feeling for the general approach adopted in the present research. Moreover, it attempted to illustrate the way in which theoretical or basic research can be useful in addressing issues of "real world" learning being currently investigated.

### General Procedure and Rationale

The research reported here has two primary purposes: first, to determine whether strong backward review effects can be obtained; and second, to examine the nature of such a review process. At the least, it is hoped that this research will begin to define the conditions under which such review effects are optimal, to identify the locus of such effects and their relationship to the probe, and then begin to sort out the possible alternative processes responsible for the effect. In addition, the results of the studies may provide insight into more general aspects of storage and retrieval of information from memory.

### Mode of Presentation

As previously observed, the review effects in the

McGaw & Grotelueschen (1972) and Rothkopf & Billington (1974) studies were small: a half-question advantage was exhibited for the matched over the unmatched criterion questions in the latter study. Backward review effects may indeed be consistently small, but on the other hand it is possible that more appropriate conditions exist which would manifest stronger effects. For example, initial learning level and proximity of the adjunct questions to the pertinent information appear to be important. Both relate to the fact that for a strong review effect to occur, information must be available in memory to be enhanced, but not so well learned that a review would be ineffective in elevating recall (especially when measured after a relatively short period of time). Such a balance might be obtained by limiting the Ss' contact with the materials, while inserting questions frequently. Thus, learning would be low enough for enhancement to occur, yet the needed information would likely be available when probed. Neither of these conditions were present in the previous two studies. The Ss controlled their own study times and the probes occurred infrequently: two questions followed every four pages of fact laden material. Not only was the review effect small, but the overall level of learning was low: in the Rothkopf & Billington experiment the criterion items had mean levels of recall at 30% or less, while the adjunct

questions (from CT2) had a mean level of only 48%. Pilot data obtained by the author appear to indicate that, in fact, a strong review effect does seem likely to occur when exposure to passages is limited in combination with frequent probing. Consequently, this procedure was followed in the present research.

In order to control the rate of exposure across Ss, the reported research dealt with prose presented orally with no opportunity for a second inspection. Probes were presented orally for Experiment I and in written form for Experiment II.

### Materials

In order to investigate and accurately describe what effects adjunct post-questions have on learning outcomes, there must be a way of specifying the content of a text and the relationship of the questions to that content. Effects can be clearly identified only if we can designate which components of the text are associated with a given question and can indicate what operations the reader or listener must perform on those components. Therefore, it is necessary to select or construct passages with known characteristics and questions with the same specificity. To satisfy these requirements, passages with an inferential structure were selected. The following example is one of five



passages used in this research (Frase, 1969).<sup>1</sup> Each has the same structure but deals with a different topic:

There are about fifteen different tribes in the country of Central Ugala. The farmers belonging to this country are peace loving, which is reflected in their art work. The hill people of Central Ugala are all farmers. The upper highlands provide excellent soil for cultivation. The outcasts of Central Ugala are all hill people. It is the custom in this country to get rid of certain types of people. The Fundalas are outcasts from the other tribes in Central Ugala.

This passage was constructed from four basic premises, establishing relationships between five classes: (A) Fundalas, (B) outcasts, (C) hill people, (D) farmers, and (E) peace loving people. The complete inferential structure can be best described as a chain in which the links or relations represent inclusion ("is a"):  $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$ . The four basic links,  $A \rightarrow B$ ,  $B \rightarrow C$ ,  $C \rightarrow D$ , and  $D \rightarrow E$ , were all explicitly stated in the passage, leaving six inclusions to be inferred. Extraneous information (filler items) was inserted between premises to make the passages appear more natural.

The Watts & Anderson (1971) study demonstrated that application questions, which necessitate more thorough processing of the text, show stronger facilitation effects than questions which require only rote recall of factual terms. Inferences well fit this category: each must be

---

<sup>1</sup>The author wishes to thank Dr. Frase for kindly making available his materials for use in this research.

verified by integrating various pieces of explicitly stated information. Consequently, the inferential quality of the passages made apparent the selection of appropriate, specifiable adjunct questions. The six inferences which can be deduced from each passage can be explicitly defined in terms of the text content and yet require more than rote recall. Hence, they were used in the present studies as the inserted questions or probes.

### Purpose of the Studies

Experiment I was designed primarily to demonstrate that performance on a subsequent recall task can be elevated by appropriately probing memory for prose passages immediately following their presentation. This enhancement is in comparison to recall under control conditions where the probes do not entail meaningful review of memory. Experiment II attempted to substantiate the reliability of this effect, as well as clarify its nature. It examined whether the facilitation could be attributed to a strengthening of the memory traces occurring at the time of the probe or to cueing of the needed information by the retrieved probe during the recall task.

## Experiment I

The focuses of the first study were to establish the existence of a review effect, identify its locus, and obtain some indication of what has been stored in memory.

### Is There an Effect?

To obtain some basic results, two variables of potential importance were controlled. First, organization was held constant across passages. Each passage maintained contiguity of links for all inferential paths (component relations needed to verify a given inference), but were presented in reverse order:  $D \rightarrow E$ ,  $C \rightarrow D$ ,  $B \rightarrow C$ , and  $A \rightarrow B$ . According to Huttenlocher (1968), this organization is second in ease of acquisition only to its reversed sequence. The second variable controlled was the number of relations required by an inferential probe for verification. Only those requiring two links for solution were employed:  $A \rightarrow C$ ,  $B \rightarrow D$ , and  $C \rightarrow E$ .

In order to use the probes as recall cues instead of as adjunct questions in Experiment II, it was necessary to first ascertain whether prior knowledge of the recall task affects performance on it. Therefore, in Experiment I Ss were given either incidental or intentional learning instructions. They all were informed of a single probe occurring after each of the five passages, but only those



receiving intentional learning instructions knew of the ensuing free recall task. A number of studies (e.g., Hyde & Jenkins, 1969; Johnston & Jenkins, 1971) report that intentional instructions do not elevate recall if the incidental task involves semantic processing. In the present research, Ss who did not know of the ensuing recall task were still required to meaningfully process the prose in order to answer the adjunct questions. It was predicted, then, that they would perform as well as the "intentional" group on the recall task. A discrepancy could have developed for two reasons. First, the intentionally instructed Ss may be able to set up an effective rehearsal scheme during the few seconds available between their response to each adjunct question and the occurrence of the next passage. Second, they may diligently attend to all of the passages in anticipation of further testing, while the incidentally instructed Ss adapt their processing in response to the demands of a single, immediate adjunct question. This could result in lower performance for the incidental group on the recall task.

A baseline reflecting memory for the passage with minimal backward review was needed. Two false probes were intended to satisfy this requirement. One involved a relation of the form  $x \rightarrow y$  (F1) where neither  $x$  nor  $y$  were elements of the passage. This probe can be immediately

rejected on the grounds of complete unfamiliarity. This means of easy rejection, especially in the incidental learning condition, should have minimized the probability that Ss reviewed the stored information, though the possibility of some kind of general scan occurring can not be denied. The second false probe was of the form  $x \rightarrow F$  (F2) where  $x$  was not an element of the passage but  $F$  was an element of the inferential chain. This can also be immediately rejected on the grounds that " $x$  is a --" is unfamiliar, but the presence of  $F$  may possibly have caused some meaningful review.

### Locus of the Effect

Given that a strong review effect can be established, the next question of interest is: Exactly what is enhanced? A careful delineation of the locus of the effect should provide information relevant to defining the process involved, whether it be an actual strengthening of memory or a cueing phenomenon occurring at recall.

Several potential patterns of facilitation existed. First, a general enhancement might have occurred; the whole passage may have benefitted by being probed. Second, only the inferential chain may be better recalled while filler information may not. Third, the enhancement may be even more selective and only recall of the links critical to

verifying the probe may be facilitated. In other words, if the probe was " $A \rightarrow C$ ?" then  $A \rightarrow B$  and  $B \rightarrow C$  would be better recalled.

Only after the locus of the effect has been determined can hypotheses be developed and tested to explain the process responsible for the review effect.

### What is Stored?

It is evident that knowledge of what is stored is crucial. Bransford and Franks (1971) used a recognition task to investigate memory for semantically related sentences. Their results indicated that Ss "spontaneously integrate" information from such sentences into wholistic ideas. Ss most confidently recognized sentences representing complete ideas even though each of the acquisition items encompassed only partial ideas. It should be noted that acquisition procedures may be responsible for this conclusion due to high levels of interference. In any case, Bransford and Franks interpreted their Ss' confidence ratings as reflecting the degree to which a sentence represented what was stored. In order to obtain similar information, a recognition task followed the free recall test in Experiment I. The 20 explicit relationships or links, 30 inferences, true filler items, and false distractors were presented to the Ss. They were instructed to judge whether



the ideas represented by the sentences were expressed in the paragraphs and, if so, whether they were explicitly stated or merely implied. The Ss rated their confidence in these answers.

It was hoped that the results would give an indication of what the Ss had stored in memory. Based on findings cited earlier suggesting that independent variables involving integration or organization affect recall (accessibility) while those involving repeated exposure or rehearsal affect recognition (availability), then an advantage is to be gained by using both types of criterion measures. Examination of both the free recall and recognition ratings of information contained in baseline paragraphs compared to those probed with true inferences should begin to provide some insight into the nature of the review process; whether a simple scan or some sort of integration occurs.

In summary, Experiment I was designed to not only demonstrate the existence of a strong backward review effect, but also to begin to provide a detailed characterization of it, at least on a descriptive level. From this basic foundation, an investigation into the actual processes responsible for the effect can be developed.

## Method

Subjects. Eighty volunteers from the introductory psychology courses at the University of Massachusetts, Amherst, were randomly assigned to one of the ten experimental conditions. The Ss were run in groups ranging in size from four to eight.

Materials. Five fictional prose passages, slightly modified versions of those devised by Frase (1969), were used. The passages were approximately 103 words long, all with the same logical structure as described earlier. Each passage dealt with a different topic: the people of a foreign country, production of new cars, a political demonstration, new astronomical discoveries, and a psychologist's speech before a PTA. The five inferential classes of each passage are given in Table 1.

The five inserted post-questions were all in the form of statements whose truth was to be verified by the Ss. They had the option of responding true, false, or ?; the third category being reserved for the case when the decision had to be made by guessing.

The three experimental probes were each true inferences constructed from two adjacent links in the inferential chain:  $A \rightarrow C$ ,  $B \rightarrow D$ , and  $C \rightarrow E$ . The two probes used to establish baseline effects were of two types:

Table 1  
Experiment I: Five Inferential Classes for Each of Five Passages

Passage	Text Points				
	A	→	B	→	C → D → E
1	Fundalas		outcasts	hill people farmers	peace loving
2	Fazollini		European cars	alloy	abuse exceptional
3	light years away		interesting	blue planets	Galaxy 4 supports life
4	Mrs. Richards		3 packs per day	chain smokers	pace floor mental disorder
5	Ronald Mertz		Chicagoan	beat students	fascist pig prosecuted



either false and completely unfamiliar ( $x \rightarrow y$ ) or false but containing a recognizable element from the passage ( $x \rightarrow F$ ).

Items for the recognition test were the 20 relationships explicitly stated in the passages, the 30 possible inferences, 10 true filler items (2 from each passage), and 30 false items presented in one of three random orders.

Design. Forty Ss were randomly assigned to each of two instructional conditions: intentional and incidental. Within each instructional condition the five probes and five passages (both Within-subjects variables) were ordered according to the same Greco-Latin square (Myers, 1966). Each square had eight replications, allowing eight Ss per condition.

Procedure. The incidental and intentional learning groups differed only in their initial instructions. The Ss were all told to listen carefully to five passages; each of which would be followed by a question testing what they had learned. Only the intentional learning group was further informed that an additional test on the material would follow the fifth passage.

The passages and probes were presented orally by a female, tape-recorded voice at a normal rate of approxi-

mately 135 words per minute. A signal immediately followed each of the five passages to warn Ss of the ensuing probe. They were given a sufficient time of 15 seconds to write their response on an answer sheet provided by the E. A second warning signal terminated the response period and immediately preceded the next passage.

Response to the fifth passage and probe was followed by free recall instructions. Each S was given a booklet consisting of five blank sheets of paper, each headed by a key word indicating the appropriate passage to recall. The order of recall was the same as in acquisition. The Ss were instructed to "write down everything you have learned from the passages in the order indicated." They were given an adequate period of  $3\frac{1}{2}$  minutes per passage for recall and were told not to refer back to a recalled passage once its allotted time was up.

After the free recall test, the Ss were told that a recognition task followed. They listened to tape recorded sentences (randomly selected from the three orders) and had 10 seconds following each to respond as follows:

Your task will be to first decide whether a sentence is true or false based on the paragraphs you heard earlier. Second, you are to rate how confident you are of that answer on a scale from one to five, where one means very low confidence and five means very high confidence. Third, if you decided the sentence was true, you must decide if it was explicitly stated or merely implied and again rate your confidence in this answer from one to five.

A 12 page response booklet was provided with eight response blocks per page. The Ss were required to circle the appropriate response in each block and were encouraged to use the full range of confidence ratings.

### Results

A number of important aspects of the data are to be addressed. First, responses to be inserted questions will be analyzed. Second, the results of the free recall criterion test will be examined. After assessing overall recall levels, primary concern will be directed towards comparing recall of passages subject to review with those which presumably were not; i.e., recall of true- versus false-probed passages. Once a general comparison has ascertained the existence of a true-probed passage recall advantage, the locus of such an effect will be sought by examining the recall of specific passage elements: links, inferences, filler items, and the probes themselves. Third, the recognition data will be investigated to determine whether or not the probing manipulation differentially affected this dependent measure.

#### Responses to the Inserted Post-questions

The five probes were scored as either correct or incorrect, with all question marks included in the second



category. Examination of these responses revealed an overall error rate of 22%. True probes were incorrectly responded to 28.8% of the time. Inspection of the Ss' recall protocols corresponding to these errors revealed that 90% of these Ss failed to report one or both of the links necessary to verify the probe.

Significantly fewer errors were recorded for false probes: 11.9% ( $t_{79} = 4.49$ ,  $p < .001$ ). Further examination of these errors showed that incorrect responses were made on 18.8% of the F2 probes, those involving one class element from the passage. On the other hand, only 5% of the F1 probes, those that were completely unfamiliar, were responded to incorrectly.

### Free Recall Criterion Test

#### Scoring of Recall Data

Each of the passages was divided into idea units; one unit represented each of the four basic links and six inferences. An additional number of units was assigned to denote filler material. The free recall protocols were then scored according to the above specifications, counting as correct exact replications of words and phrases, as well as appropriate synonyms and paraphrases. Note that the scoring was conservative in that the presence of each of the five classes was not recorded unless it was mentioned

in the context of a link or inference. One judge, blind to the experimental condition of each S, did this scoring. To establish the reliability of the scoring procedure, ten Ss' protocols from each instructional condition were randomly selected for scoring by a second judge, naive to the purposes of the experiment. Agreement as to the presence or absence of the 87 idea units in each of the twenty protocols ranged from 91% to 100%, the mean being 96.8%.

### Overall Level of Recall

Recall of the four links and six inferences per passage, averaged across all Ss and conditions was 24.2%. Calculated as a function of the serial position of the passage during acquisition, it is quite apparent from Table 2 that recall did not vary with serial position.

Generally, recall was quite high, particularly when considering how limited the Ss' exposure was to the passages. Across all Ss and conditions, a mean of 44.6% of the links, the basic structure of the passages, was recalled. In addition, a mean of 24.5% of the filler material and 10.4% of the possible inferences were reported. This latter quantity is a conservative reflection of the Ss' knowledge of the inferences. Given that a S correctly recalled two adjacent links, for instance  $A \rightarrow B$  and  $B \rightarrow C$ , there was no reason to believe he would also write

Table 2

Experiment I: Mean Percentage of Links and Inferences  
Recalled as a Function of Passage Serial Position,  
Averaged Across All Ss and Passages

Percentage Recalled	Serial Position				
	1	2	3	4	5
	24.3	24.1	25.0	25.4	21.8



down "therefore  $A \rightarrow C$ " since 1) he was not instructed to do so and most likely was trying to duplicate the actual passage, 2) the inference was obviously deducible from the reported links, and 3) he was under a time constraint. Admittedly, it is possible that at least some of the Ss were not aware of such "obvious" inferences. However, in order to obtain a less stringent estimate of the Ss' inferential knowledge, the recall protocols were rescored to include any inferences deducible from information actually recalled. Table 3 presents the revised mean percentages and shows an overall increase in inferential recall from 10.4% to 29.9%.

#### False-probed Passage Recall: A Baseline

Recall that two of the five passages heard by each S were followed by false probes; one was completely unfamiliar (F1) and the other mentioned an element from the passage (F2). The F1 and F2 error rates suggested that the hypothesis concerning the possibility that F2 probes may have engendered some review and thus differential recall (relative to F1) should be investigated. In fact, a matched t test demonstrated that recall of the four passage links from each of the two false-probed passages was almost identical ( $t_{79} = .024$ ). Since this was the case, the recall of F1 and F2 probed passages were taken as indices of

Table 3

Experiment I: Mean Percentage of Inferences  
Recalled and Implied Per S

Groups	Probe Type/passage Combination		
	True-Probed Passages	False-Probed Passages	$\bar{X}$
Incidental	32.8	27.5	30.7
Intentional	30.2	27.3	29.0
$\bar{X}$	31.5	27.4	29.9

baseline recall, unaffected by review processes, and therefore these data were combined to serve that purpose in further analyses.

### Recall: Is there an Effect?

It was initially important to determine whether Ss' recall of information from the three true-probed passages differed generally from that of the two false-probed passages. Table 4 presents data relevant to this comparison. Under both incidental and intentional instructions, the percentage of total links and inferences recalled from the passages was greater following true probes than false ( $F(1,70) = 4.82, p < .05$ ). Knowledge of the ensuing recall task did not enhance performance. In fact, the incidental group had a slight, though nonsignificant, advantage: they recalled 1.3% more of the links and inferences than the intentional group.

### Locus of the Effect

In order to determine the locus of the facilitative effect, each S's recall data were parsed into several categories. First, the six links necessary for true probe verification were identified as component relations or links:  $A \rightarrow B, B \rightarrow C$  from the  $A \rightarrow C$  probed passage,  $B \rightarrow C, C \rightarrow D$  from the  $B \rightarrow D$  probed passage, and



Table 4

Experiment I: Mean Percentage of Total Links  
and Inferences Recalled Per Ss

Groups	Probe Type		
	True Probes	False Probes	$\bar{X}$
Incidental	25.4	23.4	24.6
Intentional	24.5	21.6	23.3
$\bar{X}$	25.0	22.6	23.9

$C \rightarrow D$ ,  $D \rightarrow E$  from the  $C \rightarrow E$  probed passage. The remaining six links from the true-probed passages, those unrelated to the verification task, were designated as irrelevant links. The other five categories were as follows: 1) true-probed passage inferences, 2) true-probed passage filler items, 3) links recalled from the false-probed passages, 4) false-probed passage inferences, and 5) false-probed passage filler items.

Link recall. Consideration was first given to the status of the links. An adjustment of the scores was necessary prior to the analysis to account for preexperimental differences in recall existing between the three link categories. Data from false-probed passages indicated that, in general, the outer links of the differential chains ( $A \rightarrow B$ ,  $D \rightarrow E$ ) were less likely to be recalled than the center links ( $B \rightarrow C$ ,  $C \rightarrow D$ ). The link categories were comprised of differing quantities of these two classes, resulting in the expectation that the percentage of links recalled per category would differ even prior to any experimental manipulation. To correct for these differences; i.e., equate the preexperimental expected recall levels of the three link categories, it was sufficient to compute category scores by weighting the percentage of center and outer links recalled in each category by their relative frequency of occurrence.

Table 5 presents the adjusted mean percentages recalled for the three link categories per instruction type. Within and across the two experimental groups, component links appera to be recalled better than either irrelevant or false-probed passage links. The main effect due to link category was significant at the .05 level ( $F(2,140) = 3.09$ ). A set of four contrasts (Bonferroni  $t$ ) failed to reveal the source of the effect, though those comparing component link recall with each of the other link categories did approach significance ( $EW = .05$ ). Again, the intentional group did not perform better than the incidental.

The above analyses included data from true-probed passages whose probes were responded to incorrectly (28.8%). Under such circumstances, any prediction of enhanced recall due to a backward review clearly rests on uncertain ground. True-probe errors provide no basis for assuming that a review of memory ever took place or that information was available for review, both prerequisites for enhancement to occur. By eliminating the recall data associated with these errors, it was hoped that a clearer picture of the enhancement would emerge. By doing so, mean recall of component links rose to 57.2%, while 37.0% of the irrelevant links were recalled. The component link category then showed a 20.2% advantage over the irrelevant link



Table 5  
 Experiment I: Mean Percentage (Adjusted)  
 of Links Recalled

Group	Link Type			
	Component Links	Irrelevant Links	False-Probed Passage Links	$\bar{X}$
Incidental	52.2	47.8	44.7	48.2
Intentional	44.4	37.2	41.3	40.9
$\bar{X}$	48.3	42.5	43.0	44.6

category and a 14.2% advantage over the false-probed passage links.

Inference recall. It can be seen from Table 6 that slightly more inferences (2.3%) were recalled from true-probed passages than from false. Also, averaged across probe type, the intentional group recalled slightly more than the incidental. When the percentage recall scores were subjected to an analysis of variance, these differences did not prove to be significant. Significance was attributed only to the interaction of the third factor, probe-passage order, with both instructions ( $F(4,70) = 3.85$ ,  $p < .01$ ) and probe type ( $F(4,70) = 5.23$ ,  $p < .01$ ).

Filler recall. On the whole, Ss recalled a mean of 24.5% of the filler material regardless of the experimental manipulation involved. An analysis of variance failed to demonstrate any significant recall differences due to the main or interaction effects of probe type, instructions, or presentation order.

Probe recall. To help determine the status in memory of the true probes, their relationship with the other two-link inferences was investigated. Ss recalled a mean of 22.5% of the true probes, 14.0% of the remaining two-link inferences from the true-probed passages, and 12.7% of those from false-probed passages. No consistent differ-

Table 6  
 Experiment I: Mean Percentage of  
 Inferences Recalled Per S

Group	Probe Type		
	True Probes	False Probes	$\bar{X}$
Incidental	10.0	9.2	9.7
Intentional	12.6	8.8	11.1
$\bar{X}$	11.3	9.0	10.4



ences were evident between instruction types. An analysis of variance revealed that the main effect of inference type was significant ( $F(2,140) = 5.23, p < .01$ ), while the other main effects of instruction type and probe passage order did not reach significance. Moreover, inference type interacted with probe-passage order ( $F(8,140) = 5.76, p < .001$ ). The recall advantage enjoyed by the true probes over both of the other two-link inference types was substantiated by means of the Newman-Keuls procedure ( $\alpha = .05$ ). Recall in the latter two categories did not significantly differ.

### Recognition Criterion Test

#### Scoring of the Recognition Data

The recognition data were collected in such a way as to reflect the degree of confidence Ss had in their judgements of whether or not each test item was true based on information contained in the passages and, if judged true, whether that item had been directly stated or merely implied. These data were scored as follows: Ss' ratings were converted into numerical values. Responses of false with confidence ratings of five through one were converted to zero through four, respectively. True responses with confidence ratings of one through five were changed to five through nine, respectively. Thus a ten point scale was created ranging from a strong false (zero) to a strong true

judgement (nine). A similar scale was constructed for the stated-implied dimension with zero representing a very confident "stated" response and nine a very confident "implied" response.

### Does Probing Differentially Effect the Recognition of True Items?

The recall data demonstrated that the strength of the facilitative effect was concentrated on the component links from which the true probes were derived. Therefore, attention was first directed to a comparison of the true-false confidence ratings assigned to component versus irrelevant links. A mean difference score was calculated for these categories for each S. Component and irrelevant links were not differentially rated on the true-false scale as demonstrated by the obtained grand mean of .006. The results of an analysis of variance supported this finding and, furthermore, indicated that component and irrelevant links were not significantly affected by type of instruction, order of probe-passage combination, or their interaction.

A series of additional analyses served to substantiate the finding that true probes did not enhance the recognition of true items (links, inferences, or filler) over those same items subject to false probes. There was one exception: true probes were more confidently recognized as being true ( $\bar{X}$  = 7.00) than other two-link nonprobes from

either true-probed passages ( $\bar{X}$  = 6.38) or false-probed passages ( $\bar{X}$  = 6.42). An analysis of variance bore this out. The effect due to inference type was significant at the .01 level ( $F(2,140) = 5.24$ ), while neither instructions nor probe-passage order contributed to a main effect. In addition, inference type interacted with probe-passage order ( $F(8,140) = 3.85$ ,  $p < .001$ ). The Scheffe' procedure confirmed that probe ratings were reliably higher than either of the other two categories which did not differ from one another ( $F(2,140) = 10.4$ ,  $p < .025$ ;  $F(2,140) = .05$ ,  $p > .025$ ).

#### Recognition of True versus False Items

It was quite apparent from the data that the Ss could distinguish between true ( $\bar{X}$  = 6.79) and false items ( $\bar{X}$  = 1.79). Weighted difference scores were evaluated using an analysis of variance in order to determine if these scores varied along any other dimensions. The scores were assessed via the Greco-Latin square design (Myers, 1966, p.260). As anticipated, the scores changed as a function of probe ( $F(4,280) = 3.61$ ,  $p < .01$ ) and passage ( $F(4,280) = 20.41$ ,  $p < .001$ ) but were unaffected by sequential position, order, instructions, and all interactions. The Scheffe' procedure was used to contrast the three true-probed conditions with the two false-probed conditions. The groups differed significantly ( $F(4,280) = 12.3$ ,  $p < .05$ ), showing that Ss'



judgements of true versus false statements diverged more when the statements came from true-, as opposed to, false-probed passages.

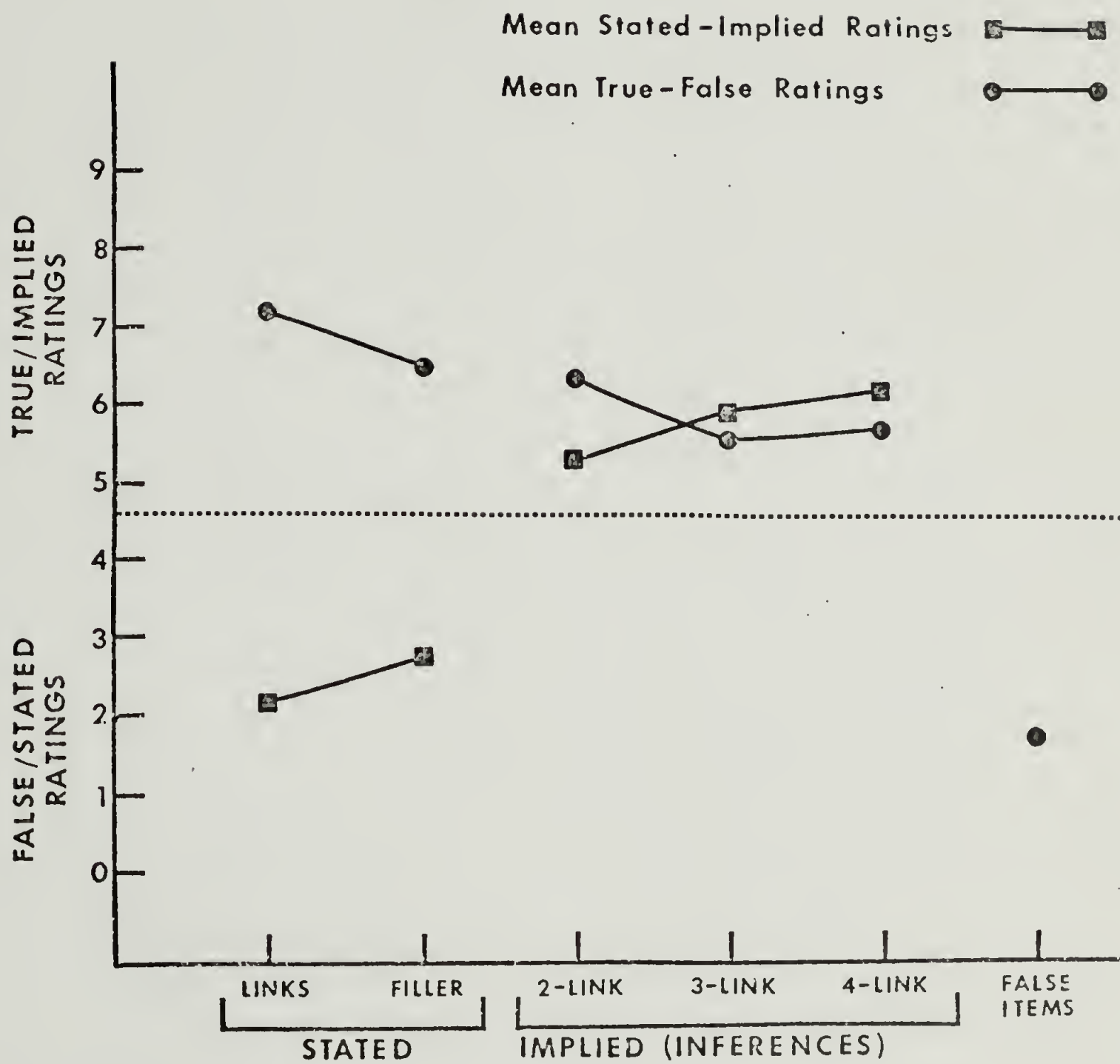
### Stated-Implied Judgements

Despite the fact that the experimental manipulations had little effect upon the recognition data, the task did demonstrate the fact that the Ss were well able to not only distinguish true from false statements but also stated from implied. It can be seen from Figure 1 just how the two judgements were related and changed as a function of item type. Confidence in the truth of items was the highest for those that were explicitly stated in the passages and appeared to decrease for inferences as the number of links needed for their derivation increased. Similarly, Ss were most confident that the stated item had, in fact, been stated while their confidence in the implied nature of the inferences increased with the number of component links.

### Discussion

The primary goal of Experiment I was achieved: the existence of a backward review effect was substantiated. The strength of such an effect was concentrated on information whose integration was necessary for the probes' verification. Though in the overall analysis this trend was statistically nonsignificant, reasonable grounds exist

Figure 1



for considering only recall data from passages whose probes were responded to correctly. When these data were appraised component link recall was 20.2% higher than for irrelevant links and 14.2% higher than for false-probed passage links. While the reliability of such a finding must necessarily be ascertained, it apparently represents a considerably stronger effect than previously demonstrated in the inserted question literature.

### Subordinate Issues

A number of points of secondary interest were raised. For instance, no evidence of a forward shaping effect was found in the experiment: recall of the passages did not improve as a function of serial position nor was recall of the first passage encountered depressed by the lack of any preceding probe. Evidence of what Rothkopf (1966) terms "direct instructive effects" was found. Ss recognized, as well as recalled, the two-link inferences used as true probes significantly better than other two-link inferences.

Experiment I further demonstrated that knowledge of the recall task at the time of acquisition did not influence recall: the incidental and intentional groups' performances were quite similar. Not only was the instruction-type variable nonsignificant as a main effect in all of the analyses, it did not interact with probe type. The lack of an interesting effect of intentionality



supports the argument that success in the inserted question task required Ss to process the passages at semantic level, regardless of other instructions. It also indicates that if, in fact, knowledge of the recall task was responsible for differential learning strategies (e.g., interpassage rehearsal, better attentiveness, etc.), then their effectiveness was negligible.

The recognition task did not contribute directly to further understanding of the review process. The various probes did not cause discernable differences to occur in recognition of the links alone or links and inferences taken as a whole. This lack of results is pertinent for at least two reasons. First, it demonstrates the importance of carefully choosing a criterion measure which will reflect effects of independent variables. Second, the fact that recall was affected while recognition was not suggests that, in line with more basic research cited earlier, the review process may involve the organization or integration of information as opposed to a simple scan. This brings up issues concerning the character of the review process which will be dealt with next.

#### The Nature of the Backward Review

A review effect has been established; hence, a next consideration is whether the facilitation is basically a storage or retrieval phenomenon. For instance, facilitation

could be the result of "directed rehearsal". In other words, attention is directed to stimuli while they are still available in memory, at which point they are rehearsed and then better recalled. Another possibility is that a strong facilitation can best be accounted for by Craik and Lockhart's (1972) notion of "depth of processing". By verifying an inferential probe while the basic premises are still available, the links are integrated, thus more deeply or thoroughly processed. Consequently, the material is better recalled. Both of these possibilities affect storage and can be considered to result in a strengthening of memory. A third potential explanation exists: the probe may become incorporated as an additional retrieval cue. By probing the S after the presentation of a passage, he is being reexposed to two classes from the passage. At recall, he may remember the probe and use the two classes to prime or regenerate information which might otherwise have been omitted. These possibilities need not be mutually exclusive. Experiment II was directed toward examining these explanations.

## Experiment II

Experiment II was designed to establish the reliability of the backward review effect and to begin to experimentally distinguish between the storage and retrieval explanations of the effect put forth in the discussion of Experiment I.

In this second study, the probes were given as inserted questions immediately following each passage ( $\overline{PC}$ ), were withheld until the time of recall when they were explicitly given as retrieval cues ( $\overline{PC}$ ), or were experienced as both inserted questions and again as cues. It was predicted that if enhanced recall of the component links can be best accounted for by a simple cueing process alone, then the explicitly cued  $\underline{S}$ s should do at least as well as  $\underline{S}$ s given only inserted questions. The cued  $\underline{S}$ s would have the advantage of not having to first recall the cues before generating the component links. In addition, the  $\overline{PC}$  group would have the complete 15 second interpassage interval to set up a rehearsal scheme. On the other hand, if a change in storage is an important determinant of the phenomenon, whether or not cueing also plays a role, the  $\overline{PC}$  group should perform better than the  $\overline{PC}$  group. This prediction rests on the assumption that the probes are responsible for somehow strengthening memory for the component relations by initiating additional processing



before the links' availability and/or accessibility suffer major losses. Possible ways in which this might occur were discussed earlier. This prediction further assumes that any interpassage rehearsal by  $\overline{PC}$   $\underline{S}$ s is not as effective as the probing technique in elevating recall. It was hoped, then, that comparisons between the  $\overline{PC}$ ,  $PC$ , and  $\overline{PC}$  groups would provide information that would begin to help differentiate the storage and retrieval hypotheses, as well as compare the effectiveness of the probing technique with  $\underline{S}$ -provided rehearsal schemes.

Those  $\underline{S}$ s not encountering inserted questions must necessarily be informed of the recall task since it is likely that  $\underline{S}$ s would surmise the existence of some sort of test anyway. Since type of instructions did not affect performance in Experiment I, this did not create a problem. The intentionality variable was eliminated in Experiment II: all  $\underline{S}$ s were informed of the post test.

The results of Experiment I suggested several other procedural improvements. Of the  $\underline{S}$ s making errors on true probes in Experiment I (28.8%), 90% failed to report one or both of the critical links. One probable explanation is that the missing link or links were never stored due to inattention or lack of sufficient processing etc., thus explaining the probe error and ensuing lack of recall. In this case, facilitation could not occur. Review or addi-

tional processing presupposes storage and it is this added involvement with the information upon which the enhancement argument rests. Alternatively, the component links may have been stored but were simply unavailable for review at the time of the probe, again accounting for the probe error. Since the availability problem can be temporary, this option helps to explain the occurrence, though infrequent, of probe errors occurring along with later recall of the component links. Of course, if the probe was misheard or otherwise misinterpreted, this could also explain the above phenomenon. It is clear, in any case, that since probe errors accompanied poor recall of critical links, lower error rates are desirable. In view of this, the probes were presented in written form in Experiment II. It was hoped that this change in procedure would help to alleviate errors attributable to probe processing and decrease the short term memory load by making the probes available for reinspection. F1 and F2 probes induced different error rates, but not recall, in Experiment I; thus, to improve error rates, the F2 types were abrogated in Experiment II in favor of two completely false probes (F1) per S.

Since the recognition task in Experiment I did not provide any substantial information about the process of interest, it was not used in Experiment II. Another change

was made in ordering the passages for acquisition; all Ss received the passages in the same order. Presentation order was intentionally confounded with passages so that data for each passage could be looked at individually without having to consider order effects.

In conclusion, the position of the probes in the prose learning task of Experiment I was manipulated in order to gain knowledge as to what role the inserted questions play in the enhancement of the passages' later recall.

#### Method

Subjects. 120 volunteers from introductory psychology courses served as Ss. They were run in groups ranging in size from four to fifteen. Ss were randomly assigned to conditions.

Materials. The five passages and probes in Experiment I were again used with one exception; both false probes were completely unfamiliar.

Design. The within-subjects variable, cue-probes, was ordered according to a 5x5 Latin square. Three identical squares were used for the levels of the between-subjects variable, probe-cue position ( $\overline{PC}$ ,  $PC$ ,  $\overline{PC}$ ). The passage variable (within-subjects) was intentionally confounded with column effects in the square, thus all Ss were administered



the passages in the same order. The squares were replicated eight times, allowing eight Ss per ordering.

Procedure. All Ss were instructed to listen carefully to the five passages, which would be followed by a test to determine what they had learned.

The procedure governing the presentation of the passages and the free recall test were the same as in the first study, with the following exceptions. A 15 second blank period followed each passage. Those Ss in probe conditions were given this time to read the appropriate probe from a provided booklet and respond to it, while those in the cue-only condition were given no instructions as to how to employ this time. For the free recall test, the Ss were told that each response sheet was headed by a key word or phrase indicating the passage to be recalled. The key words for all the passages in the  $\overline{PC}$  condition and for the two F1/cued passages in the PC and  $\overline{PC}$  conditions were the same as in Experiment I. The key words for the other three passages in the PC and  $\overline{PC}$  conditions were the two-link inferences used as true probes.

## Results

### Responses to the Inserted-questions

The probes were scored as in Experiment I. The overall error rate for the inserted questions was 16.8%,

significantly lower than in Experiment I ( $t_{79} = 1.90$ ,  $p < .05$ ,  $\bar{D} = 5.2\%$ ). True probe inaccuracies were recorded 26.3% of the time. Of the  $S$ s making these errors, 87.3% later failed to recall one or both of the component relations corresponding to these probes. The false probe error rate was 2.5%, substantially lower than that for the true probes.

### Free Recall Criterion Test

#### Scoring of the Recall Data

Scoring of the free recall protocols was done as in Experiment I. The reliability of the scoring procedure was again established. The agreement of a second judge on 30 randomly selected protocols ranged from 88.3% to 100%, with a mean of 95.6%.

#### Overall Recall

As in Experiment I, overall recall was high given the way in which the material to be learned was disseminated. Excluding the three probes,  $S$ s on the average recalled 22.6% of the possible links and inferences (23.6% for  $PC$  only, including probes). Recall as a function of serial position was not examined since serial position was confounded with passage order.

Broken down by category, mean recall of links was 43%, inferences minus probes was 7.4%, and filler items was 22.4%. The same figures for  $PC$   $S$ s only, directly compar-

able to those of Experiment I, are 46.6%, 8.8%, (including probes) and 21.6%, respectively.

When inclusion of inferences implied by recalled information was considered, recall of inferences rose from 7.4% to 20.9% (see Table 7).

### Recall: Is There an Effect?

Table 8 shows the mean percentage of total links and inferences recalled by the three probe-cue groups from true-probed/cued passages and from false-probed/not cued passages. Since the cued groups were given three of the inferences as retrieval cues, recall of these items was not considered. Averaged across the experimental groups, recall of passages which had previously been probed and/or cued with the two-link inferences exceeded the recall of passages which had been probed with irrelevant statements and/or had not been cued ( $\bar{D} = 4.7\%$ , Scheffe' test:  $F(4,420) = 17.49$ ,  $p < .05$ ). The data suggest that this advantage was more pronounced for the two groups experiencing probes than for the cue-only group: a 6.1% difference compared to 1.9%. This interaction, however, was not significant; nor was the probe-cue condition main effect, though a trend favoring higher recall levels for probed groups seemed to exist. There were only small, nonsignificant differences evident between the recall of false-probed passages and those that had been neither probed nor cued. This lends support to



Table 7  
 Experiment II: Mean Percentage of Inferences  
 Recalled and Implied Per S

Group	Probe/Cue Type		
	True Probed and/or Cued	False Probed and/or Not Cued	$\bar{X}$
$\overline{PC}$	20.8 <sup>a</sup>	22.6	21.5
PC	25.5	20.0	23.3
$\overline{PC}$	17.6	18.5	18.0
$\bar{X}$	21.3	20.4	20.9

<sup>a</sup>Including recall of the probes: 26.0%; 51.8% of probes recalled and implied

Table 8  
 Experiment II: Mean Percentage of Total Links  
 and Inferences Recalled Per S

Group	Probe/Cue Type		
	True Probed and/or cued	False Probed and/or Not Cued	$\bar{X}$
$\overline{PC}$	25.9 <sup>a</sup>	20.7	23.8
PC	26.5	19.6	23.7
$\overline{PC}$	21.0	19.1	20.3
$\bar{X}$	24.5	19.8	22.6

<sup>a</sup> 22.5% of probes recalled

the assumption that false probes do not initiate the same kind of review of passage material as do true probes. The ability of the dependent variable to detect any possible effects due to memory scans made in response to false probes must be considered.

The effects of the various probes on individual passages was investigated and is illustrated in Figure 2. No consistent patterns were evident to suggest that the various true probe types were differentially but systematically effective in facilitating recall.

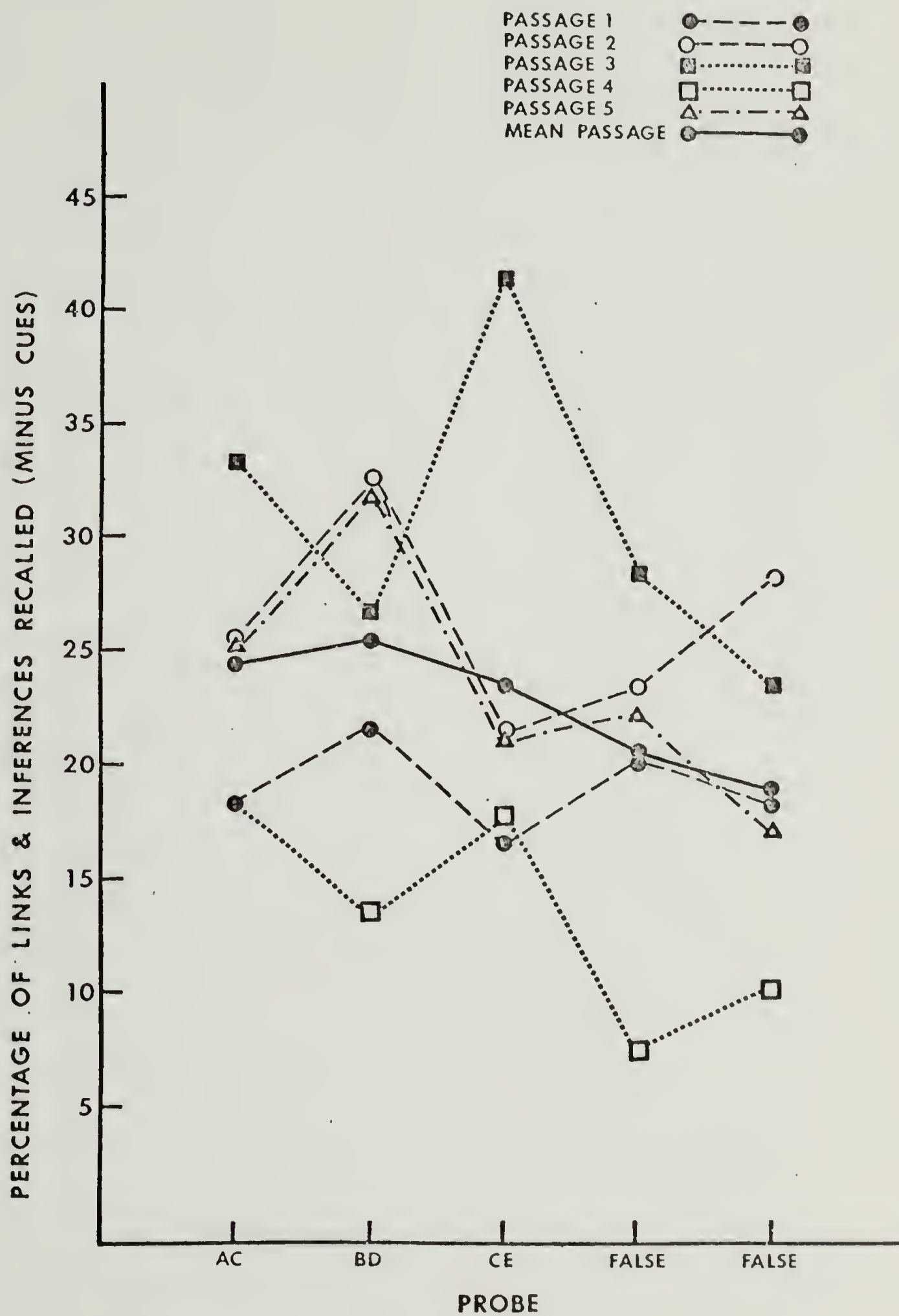
The  $\overline{PC}$  group represents a replication of Experiment I. Important aspects of its data, including recall of the probes, will be presented where appropriate. The  $\overline{PC}$  group recalled 25.6% of the possible links and inferences from the true-probed passages. For the  $\overline{PC}$  group then, a mean difference of 4.9% was obtained between true- and false-probed passages, only slightly higher than the 2.4% of Experiment I.

### Locus of the Effect

As in Experiment I, the recall data were parsed into various categories. Recalled links, inferences, and filler items were separated into six groups according to whether they came from true- or false-probed/cued passages. Links from true-probed/cued passages were further divided into component and irrelevant links, depending upon whether or



Figure 2



not the relations were germane to the derivation of the probe inferences.

Link recall. Before the link recall data were examined,  $\bar{S}$ s' scores were adjusted as in Experiment I. The key data of Experiment II are presented in Table 9. Collapsed over groups, the indication from Experiment I that component-link recall was enhanced over that of both irrelevant and false-probed passage links was confirmed. The link category main effect was significant ( $F(2,210) = 7.90$ ,  $p < .01$ ), as was the component versus irrelevant link contrast (Bonferroni  $t_{210} = 3.30$ ,  $p < .01$ ). No meaningful differences were evident between irrelevant and false-probed passage link recall ( $t_{210} = .19$ ). Comparing overall group performance, there is some indication that  $\bar{S}$ s who were not probed remembered less than those who were; however, this effect was not significant. Of primary interest is the interaction of probe-cue conditions with the link categories ( $F(4,210) = 3.39$ ,  $p < .01$ ). This effect appears to be due to the recall of component links by those groups who experienced probes. That it is only these two cells that exhibit superior recall is evidenced by a series of analyses. The component versus irrelevant link contrast was shown to be statistically the same for the two probed groups ( $\bar{D}_{c-i} = 15.1\%$ ,  $t_{105} = .95$ ), while their mean contrast was significantly larger than that for the cued only group

Table 9  
 Experiment II: Mean Percentage (Adjusted)  
 of Links Recalled Per S

Group	Link Type			
	Component Links	Irrelevant Links	False-Probed Passage Links	$\bar{X}$
$\overline{PC}$	57.5	39.1	43.4	46.7
PC	53.1	41.3	39.1	44.5
$\overline{PC}$	37.8	41.6	37.8	39.1
$\bar{X}$	49.5	40.6	40.1	43.4



( $D_{c-i} = -3.6\%$ ,  $t_{105} = 3.15$ ). Disregarding the two cells of interest, the largest remaining cell mean did not differ significantly from either the smallest between- or within-subjects cell: the  $t$  statistics did not reach significance despite the inflation of the experiment-wise error rate.

To summarize, recall of links in all categories under all conditions was equivalent except for those links necessary for true-probe verification. These component links enjoyed a 15.1% and 14.1% recall advantage over irrelevant links and false-probed passage links, respectively. When recall data from true-probed passages whose probes were responded to incorrectly were eliminated, the former advantage rose to a substantial 24.8%. Such an adjustment was accounted for by an increase in component link recall: its mean rose from 55.3% to 63.9% while the irrelevant-link mean remained virtually the same (40.2% versus 39.1%).

Inference recall. The data are summarized in Table 10. Looking at the percentage of inferences (minus true probes) recalled following false probes and/or no cues compared with true probes and/or cues, there is a suggestion that fewer items are recalled in the former conditions. A slight overall recall advantage seems to be present for the probed groups. As in Experiment I, these trends did not prove to be significant.

Again, in order to compare the results directly with

Table 10  
 Experiment II: Mean Percentage of  
 Inferences Recalled Per S

Group	Probe/Cue Type		
	True Probed and/or cued	False Probed Not and/or Cued	$\bar{X}$
$\overline{PC}$	7.8 <sup>a</sup>	6.2	7.1
PC	9.7	6.9	8.5
$\overline{PC}$	6.8	6.6	6.7
$\bar{X}$	8.1	6.6	7.4

<sup>a</sup> 22.5% of probes recalled

Experiment I, data for the probe-only group were reexamined, including recall of the three probes. Ss remembered more inferences from true-probed passages,  $\bar{X} = 10.3\%$ , than false-probed passages,  $\bar{X} = 6.3\%$  ( $F(1,35) = 10.50$ ,  $p < .05$ ), the advantage being accounted for by enhanced recall of the true probes. This pattern of results was found in Experiment I also, but was not significant.

Filler recall. A mean of 22.4% of the filler items was recalled by the Ss. There were no indications of differential recall as a function of either probe-cue condition or probe/cue type, substantiating the same finding from Experiment I.

Probe recall. Recall of the probes by Ss in the  $\overline{PC}$  condition was considered in relation to that of the other two-link inferences. As in Experiment I, probes ( $\bar{X} = 22.5\%$ ) were better recalled than either nonprobes from true-probed passages ( $\bar{X} = 12.9\%$ ) or from false-probed passages ( $\bar{X} = 11.3\%$ ), whose recall did not differ. This was confirmed following a significant main effect due to inference type ( $F(2,70) = 4.23$ ,  $p < .025$ ), by use of the Neuman-Keuls technique comparing the three means ( $\alpha = .05$ ).

## Discussion

Experiment II reaffirmed the existence of a backward review effect. As was first indicated in Experiment I, a breakdown of the data revealed the locus of the review effect to be concentrated on the links. Specifically, component links were better recalled than either irrelevant links or links from false-probed passages (recall in the latter two categories did not differ). This advantage was evident only for  $\underline{S}$ s who experienced probes. These results yield an important implication concerning the nature of the facilitative effect. Clearly, the enhancement cannot be attributed to a simple cueing process as described in the introduction; when given the corresponding inference,  $\overline{PC}$   $\underline{S}$ s were no more successful in generating the component links than irrelevant or non-cued passage links.

The evidence is not sufficient, however, to entirely discount the possibility that the probes function, at least in part, as retrieval cues. Two points are to be made. First, the procedures employed in the present research may not be a fair test of the cueing hypothesis. Each cue correctly related two class elements by way of inference. The class elements had previously been explicitly stated only within the context of other relationships. Evidence does exist (the "encoding specificity hypothesis": Tulving & Thomson, 1970; and Thomson & Tulving, 1973) which



suggests that such contextual changes may be responsible for the cues' ineffectiveness. This issue can be addressed in future research by cueing with single class elements; i.e., "Write down what you can remember about the passage involving A and C (instead of A is a C)."

A second point to be made centers around the notion of availability and accessibility. That is, in the  $\bar{P}C$  condition the cues may have allowed access to as much information as was available for recall. The overall analysis of the percentage of links and inferences recalled does suggest that some small benefit was gained by the cueing procedure. The probes may also have served as retrieval cues at the time of recall, but only subsequent to their more important role of improving the status in memory of the critical links. The exact nature of this strengthening process, whether it be through integration (deeper processing or organization) or repeated exposure (rehearsal) is yet to be investigated.

Several other aspects of the data are of secondary interest. Not only did the  $\bar{P}C$   $\underline{S}$ s derive little benefit from the cueing manipulation, but furthermore they were unable to establish an effective rehearsal scheme during the interpassage intervals (relative to probed conditions and as detected by the criterion measure). This is evidenced by the finding that there were no significant

differences between any of the false-probed passage measures and the appropriate cued measures. In particular, false-probed passage links were, if anything, recalled slightly better (3.5%--nonsignificant) than the corresponding "false-cued" links. Note that the recall of "false-cued" passages represents an intentional learning control condition (no probes, no cues).

The fact that recall did not vary across the false-probed/cued conditions also suggests that there was no noticeable forward effect operating in this experiment, at least as reflected by the free recall measure. Serial position effects could not be assessed due to a confounding with passage order.

Switching from oral to written probes did not appreciably lower the true probe error rate. However, eliminating the F2 probe in favor of two F1 probes appeared to depress false probe errors (11.9% to 2.5%), accounting for much of the overall drop from 22.0% to 16.8%. True-probed error rates in the two experiments remained quite similar. Again, 87.3% of the recall protocols corresponding to the true probe errors were missing one or both of the critical component links. Since a strong backward review effect (i.e., enhanced recall of the component links) depends upon accurate response to the inferential probes, future investigations must be directed towards identifying conditions

assuring such success.

### General Discussion

Clearly, these two studies have demonstrated a potent procedure for enhancing memory for prose. The locus of the effect was identified and an initial attempt was made to explain it in terms of processing notions; i.e., a backward review effect.

It might be argued that this review effect is just another manifestation of what Rothkopf (1966) termed direct instructive effects; that is, enhanced memory for question-relevant information. Based upon this definition, refutation is not possible or even desirable. What needs to be broadened and clarified is the scope of the term "question-relevant" and the processes implied by it. Precisely why and how is performance elevated on criterion items which have previously been experienced as inserted questions? Because of the often superficial quality of the adjunct questions and particularly because of the conspicuous disregard for internal processes, the present research suggests that the source of a substantial means of directing and strengthening memory for prose has previously been masked. Backward review is, therefore, thought to be an important effect in need of detailed investigation.

The locus of the backward review effect was limited to



links within the logical span of the inferential probes. Empirical evidence seems to rule out cueing as the sole, or major source of the effect, pointing instead to a strengthening of memory in terms of deeper processing or directed rehearsal. A recent study by Hayes-Roth and Hayes-Roth (1975) has used reaction time measures to investigate the effects of probing inferential knowledge structures. Among their results, they found that verifying an inference facilitated immediately succeeding verification of a component relation and also that the probing manipulation could create a strong direct link between class elements which were formerly connected indirectly. In agreement with the present study, it was concluded that "memorial networks are plastic and adaptive to the individual's experiences of storage and retrieval demands (p.508)." This study lends strong support to the notion of a backward review effect and incorporates both strengthening and integration effects. Exactly what role these processes play in the facilitation of prose remains to be resolved in future research.

These studies cannot be taken as definitive in their specification of the actual scope of the review. While the free recall data seem to suggest that the Ss are directly accessing the relevant information, a more global review may well be occurring. Any effects due to such a process



may not have been detected because of insensitivity on the part of the dependent measure. A number of studies (e.g., Hogan & Kintsch, 1971; and Woodward, Bjork, and Jongeward, 1973) do indicate that recall measures detect the manipulation of experimental variables affecting organizational activities, but not those affecting simple rehearsal; i.e., increased exposure to the stimuli. A process falling within this latter category may be responsible for originally locating the component links. Such a scan, undetected by the recall measure, may also occur in response to false probes. This might account for why the recognition measure in Experiment I failed to differentiate between true- and false-probed conditions. Evidence exists which further substantiates the possibility of some kind of general scan. Monk and Kintsch (1974) showed that reaction time to perform a variety of true-false tasks using prose materials was linearly related to passage length. In addition, McGaw and Grotelueschen (1972) and Rothkopf and Billington (1974) reported that facts which shared physical and semantic proximity with information tapped by post-questions were also enhanced. Further studies are obviously needed to clarify the nature of the scan and its effects on memory.

The paradigm developed here leaves many general, as well as specific, questions open for investigation.

Besides explicating the nature of the memory scan, more global issues center around how to distinguish between depth of processing and strength as brought about by repetition and also around how inferences are stored and represented in memory. Attention needs to be paid to more paradigm-bound problems, too. For instance, can this effect, as now localized, be broadened by increasing the number of component links or span of the probes? The logical and physical order of the links were confounded within the passages. Of what importance are these variables? What are the long term effects of the probing manipulation? What happens to the effect when initial exposure to the passages is varied? What variables affect true probe errors and how might they be minimized? The findings also need to be generalized to other materials and probe types. A rich future of investigation awaits.

This discussion has made evident both the theoretical and practical value of such a line of research. By relating basic findings and a processing approach associated with theoretical work to a complex applied problem, progress was made in both spheres. It is hoped then that this research has been successful, at least in some small way, in demonstrating the benefits to be gained by more interplay between basic and applied research.

## References

- Bransford, J.D. & Franks, J.J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.
- Craik, F.I. & Lockhart, R.S. Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.
- Frase, L.T. Learning from prose material: Length of passage, knowledge of results, and position of questions. Journal of Educational Psychology, 1967, 58, 266-272.
- Frase, L.T. Some unpredicted effects of different questions upon learning from connected discourse. Journal of Educational Psychology, 1968a, 59, 197-201.
- Frase, L.T. Questions as aids to reading: some research and theory. American Educational Research Journal, 1968b, 5, 319-332.
- Frase, L.T. Some data concerning the mathemagenic hypothesis. American Educational Research Journal, 1968c, 5, 181-189.
- Frase, L.T. Structural analysis of the knowledge that results from thinking about a text. Journal of Educational Psychology, 1969, 60, 1-16.
- Hayes-Roth, B. & Hayes-Roth, F. Plasticity in memorial networks. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 506-522.
- Hogan, R.M. & Kintsch, W. Differential effects of study and test trials on long-term recognition and recall. Journal of Verbal Learning and Verbal Behavior, 1971,



10, 562-567.

Hyde, T.S. & Jenkins, J.J. The differential effects of incidental tasks on the organization of recall of a list of highly associated words. Journal of Experimental Psychology, 1969, 82, 472-481.

Hyde, T.S. & Jenkins, J.J. Recall for words as a function of semantic, graphic and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 471-480.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75, 550-560.

Johnston, C.D. & Jenkins, J.J. Two more incidental tasks that differentially affect associative clustering in recall. Journal of Experimental Psychology, 1971, 89, 92-95.

Kintsch, W. Learning, Memory and Conceptual Processes. New York: Wiley, 1970.

Ladas, H. The mathemagenic effects of factual review questions on the learning of incidental information: A critical review. Review of Educational Research, 1973, 43, 71-82.

McCormack, P.D. Recognition memory: How complex a retrieval system? Canadian Journal of Psychology, 1972, 26, 19-41.

McGaw, B. & Grotelueschen, A. Directions of the effect of questions in prose materials. Journal of Educational



Psychology, 1972, 63, 580-588.

Mistler-Lachman, J.L. Depth of comprehension and sentence memory. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 98-106.

Monk, D. & Kintsch, W. Memory search I: paragraph memory and the retrieval of information. In W. Kintsch, The Representation of Meaning in Memory, New York: John Wiley & Sons, 1974, p.p. 177-193.

Myers, Jerome L. Fundamentals of Experimental Design. Boston: Allyn & Bacon, 1966.

Rickards, J.P. & DiVesta F.J. Type and frequency of questions in processing textual materials. Journal of Educational Psychology, 1974, 66, 354-362.

Rothkopf, E.Z. Learning from written materials: an exploration of the control of inspection behaviors by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Rothkopf, E.Z. Structural text features and the control of processes in learning from written materials. In: J. Carroll and R. Freedle (Eds.), Language Comprehension and the Acquisition of Knowledge. Washington D.C.: V.H. Winston and Sons, 1972, pp. 315-335.

Rothkopf, E.Z. & Bisbicos, E. Selective facilitative effects of interspersed questions on learning from written material. Journal of Educational Psychology, 1967, 58, 56-61.

- Rothkopf, E.Z. & Billington, M.J. Indirect review effects produced by adjunct questions: a replication study. Journal of Educational Psychology, 1974, 66, 669-679.
- Till, R.E. and Jenkins, J.J. The effects of cued orienting tasks on the free recall of words, 1973, 12, 489-498.
- Thomson, D.M. & Tulving, E. Associative encoding and retrieval: Weak and strong cues. Journal of Experimental Psychology, 1970, 86, 255-262.
- Tulving, E. & Thomson, D.M. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973, 80, 352-373.
- Watts G.H. & Anderson, R.C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.
- Woodward, A.E., Jr., Bjork, R.A. & Jongeward, R.H., Jr. Recall and Recognition as a function of primary rehearsal. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 608-617.

## Appendix A

## Passage 1: Cars

Recently, Auto Digest has reviewed the qualities of several cars. Any car that can stand up under continuous abuse is exceptional, says the Digest. The article also points out that cars made from space age alloys are, without exception, able to stand up under continuous abuse. All new European cars, Digest points out, are made from these advanced space age alloys. The article goes on to mention some specific factors such as weight, braking ability, and ease of steering. A major portion of the article deals with the quality of American cars in contrast to European cars. The Fazollini, a new European car produced in Italy, is one of the many cars discussed.

## Passage 2: Ugala

There are about fifteen different tribes in the country of Central Ugala. The farmers belonging to this country are peace loving, which is reflected in their art work. The hill people of Central Ugala are all farmers. The upper highlands provide excellent soil for cultivation. The outcasts of Central Ugala are all hill people. It is the custom in this country to get rid of certain types of people. The Fundalas are outcasts from the other tribes in Central Ugala.



### Passage 3: Habit Speech

Dr. Shaw, a psychologist, spoke last week at a local PTA supper in Elk River, Iowa. He discussed various habits and how they can be broken. Shaw mentioned the radical opinions that people who pace the floor at night are suffering from a mental disorder and that stuttering is just a bad habit. He also reported that all chain smokers pace the floor at night and overeat.

After the talk, parents asked him several questions about their own problems. In answer to one question, Shaw stated that anyone who smokes three packs of cigarettes a day is a chain smoker. Mrs. Richards became irritated and flew into a rage. Smoking was a dirty habit, said Mrs. Richards but that nevertheless she smoked three packs of cigarettes a day.

### Passage 4: Political Demonstration

A radical newspaper has lamented the demonstrations in Chicago during the Democratic Convention. The article made the point that the growth of Fascism is a threat to our country, and that fascist "pigs" should be prosecuted as enemies of America. The paper stated that anyone who allows the beating of unarmed students is a fascist pig.

Persuing the discussion further, a panel of reporters said that all the people of Chicago had allowed the beating of unarmed students during the demonstrations. The name



of Roland Mertz, a prominent Chicagoan, was mentioned several times by the reporters.

#### Passage 5: Astronomy

Recently, some new facts have been discovered about our universe. According to all present evidence, all planets in Galaxy IV are capable of supporting life. There is also a general agreement on the fact that all the "blue" planets are in Galaxy IV. The only interesting planets are these "blue" planets, however, later discoveries may change this judgement. Scientists consider all known planets within fifteen light years distance to be interesting. It is hoped that new discoveries will be made in the near future which will extend this knowledge.

## Appendix B

### Experiment I: Passage Instructions

You are going to listen to a recording of five short fictitious passages, each on a different topic. After each passage you will be given fifteen seconds to judge a statement true or false based solely on the content of the passage you just heard. Once you have varified whether it is valid or not, enter the appropriate answer, true or false, on the answer sheet provided. If you have no ides whether the statement is valid or not, do not guess, simply put a question mark in the answer space. The statements will be presented on the tape. Following each passage, you will hear a click followed by the statement and the fifteen second answer period. A second click will immediately precede the next passage. (For the Intentional Learning Condition only, the following sentence was now included: Following the passages you will be further tested on what you have learned.) Are there any questions?

### Experiment I & II: Recall Instructions

Now I would like you to write down everything that you learned from the five passages you have just heard. I will tell you when to begin. There is a page for each passage in the booklet before you. If you need more room, use the back of the sheet. The topics for each passage are given at the head of each page. Do them in the order in which

they are presented, and do not turn back once you have completed a topic. You will have  $3\frac{1}{2}$  minutes to recall each passage. I will signal when you are to begin and when to start each additional passage. Are there any questions? Begin.

### Experiment I: Recognition Instructions

One last task is involved, a recognition test. You will hear a number of recorded sentences, after each of which you will be given ten seconds to respond as follows: Your task is to first decide whether the sentence is true or false based on the paragraphs you heard earlier. Second, you are to rate how confident you are of that answer on a scale from one to five where one means very low confidence and five means very high confidence. Third, if you decided the sentence was true, you must decide if it was explicitly stated or merely implied and again rate your confidence in this answer from one to five.

Circle your responses on the answer sheet, noting that the sentences are numbered as are the answer spaces. Please use the full range of confidence ratings. Are there any questions?

### Experiment II: Passage Instructions--Probed Conditions

You are going to listen to a recording of five short fictitious passages, each on a different topic. After each passage you will be given fifteen seconds to judge

a statement true or false based solely on the content of the passage you just heard. Once you have varified whether it is valid or not, enter the appropriate answer, true or false, on the answer sheet provided. If you have no idea whether the statement is valid or not, do not guess; simply put a question mark in the answer space. The statements are in the booklet face down in front of you. Following the first passage you will hear a click. Turn the booklet over and respond to the first question. After the fifteen second answer period is up, you will hear a second click followed by the next passage. Remember, do not turn the booklet page to the next question until you hear the click following the appropriate passage. After you have completed this task, you will be further tested on what you have learned. Are there any questions? We will now begin.

#### Experiment II: Passage Instructions--Nonprobed Condition

You are going to listen to a recording of five short fictitious passages, each on a different topic. Following each passage you will hear a click, after which a fifteen second rest period will occur. A second click will immediately preceed the next passage. After you have listened to all five passages, you will be tested on what you have learned. Are there any questions? We will now begin.



