

A Comparison of Texts and Their Summaries: Memorial Consequences

LYNNE M. REDER AND JOHN R. ANDERSON

Carnegie-Mellon University

Chapters from college textbooks in diverse fields were compared with summaries constructed to convey the main points. A series of studies demonstrate consistent advantages for summaries. Summaries maintained their advantages at retention intervals of 20 minutes, 1 week, and 6 to 12 months. Summaries were superior both for questions directly taken from the text and for inference questions that required the subject to combine facts that had been studied. A transfer task looked at ability to learn new, related material as a function of how the previous material was learned. Summaries yielded better transfer. Reaction time differences showed the same pattern as percentage correct. Summaries maintained their superiority even when the main points in the text were underlined.

A typical college text contains about 400 pages of text and about 150,000 words. No educator seriously believes that a student will commit all 150,000 words to memory verbatim. What then, do the author and teacher expect the student to remember? In part, the text is intended to communicate a certain set of skills for reasoning or thinking cogently within that field. However, another important function of the text, judging from typical examination questions, is to communicate facts.

If one counts the number of propositions in a textbook, most modern texts contain tens of thousands of facts. No teacher can seriously expect the student to retain all of these facts. A number of cognitive scientists have proposed theories of the structure of text (e.g., Crothers, 1972; Frederiksen, 1975; Kintsch, 1974; Kintsch & van Dijk, 1975; Mandler & Johnson, 1977; Meyer, 1975; Rumelhart, 1975; Stein & Glenn, in press; Thorndyke, 1977) and these analyses imply which facts are more likely to be re-

tained. The predominant analysis of text has structured the propositions or idea units hierarchically where the more central or important propositions are represented higher in the hierarchy. Investigations of these representations have found that propositions higher and more central in these hierarchies are better recalled, more accurately recognized, and more rapidly verified (e.g., Kintsch, 1974; Kintsch & van Dijk, 1975; Meyer, 1975).

These results suggest that memory for text nicely meshes with a writer's communicative intent. The main points are best remembered and these are the propositions that the writer most wants remembered. One might wonder whether this coincidence is accidental or causal. It might be that the reader implicitly recognizes the importance of the central points and assigns greater capacity to their processing. Or it might be that main points, due to their position in the logical structure within a text, are better remembered even when amount of processing time is controlled. Meyer and McConkie (1973) report that centrality in a hierarchy is a better predictor of recall than a subject's rating of importance of propositions. This suggests that better memory for certain propositions is due to logical structure and not a subject's perception of importance.

This research was supported by ONR contract N00014-78-C-0725. Correspondence concerning this paper should be sent to Lynne M. Reder, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pa. 15213. We wish to thank C. Grimaldi, S. Matteo, and especially H. Wall for assistance in material construction, collection and analysis of data.

Function of Details

From the notions presented thus far, it is difficult to make a case for the inclusion of details in texts. Although the hierarchical text theories have not explicitly committed themselves on this issue, one would think that the detailed propositions under a main point would not have a facilitative effect in memory for prose.¹ Assuming such hierarchies are searched in a top-down manner, access to the details depends on getting the main points but there would be no dependence of main points on details. If anything, details should take study time away from the main points. Also, to the extent that subjects have difficulty in identifying the main points in a text, inclusion of details would tend to interfere with the extraction of the central ideas. Bartlett (1978) reports that students do have difficulties extracting the main points of the text and that their recall suffers as a consequence.

Of course, there are probably many arguments that can be made for why students read the long, detailed version of a text rather than a summary. For one thing, while the writer cannot seriously expect the student to retain all of the text, some of the details will probably be retained and provide illustrations and expansions of the main points.

Another goal of the inclusion of details is to acquaint the student with the argument structure of the field. Even if a student does not remember specific arguments, he or she

¹ Kintsch (1974) and Kintsch and van Dijk (1978) propose that important propositions are remembered better than less important ones because the important ones are *reinstated* during the processing of the lesser propositions. This position might be considered an argument for the importance of details in a text. However, the level of analysis in their theory is much finer than that in the current discussion. The reinstatements they refer to tend to be within the same sentence. For example, the first proposition that uses the shared argument is considered to be the superordinate of the proposition that contains the repeated argument. In contrast to this level of analysis we are considering as details whole sentences or even a paragraph that serves to illustrate a main point contained in another sentence.

may still be able to abstract the type of argument used and thereby regenerate it at some future time.

These potential functions of details, however, do not address the issue of whether details support memory for the main points of a text. If the principal educational function of a text were to communicate the central facts, would the author be justified in including details? One argument that can be made is that embellishments allow the reconstruction of the main points. That is, the details imply the main points although the converse is not true. This position is in implicit opposition to the hierarchical theories as described above. If one views memory for a text as a network structure rather than as a strict hierarchy of propositions, recall of details allows access to the main points to which those details are connected. When the main points are forgotten, one can still infer or reconstruct the main points from the details that imply them. Consider the following three sentences that are found in one of the introductory chapters we used in some of our experiments. First, there is a general statement:

(1) Africa has had less rapid development in the 1960's than the 1950's. Second, we have two statements at a level of detail below this general statement:

(2) There were decreases in administrative efficiency associated with the move from colonial to independent status in many countries.

(3) In the 1960's, there was an unwise emphasis on manufacturing at the expense of agriculture.

Sentences (2) and (3) imply (1), but not vice versa. The position that details support memory of the main points can be dubbed *the redundancy hypothesis*.

To summarize, listed below are the arguments for *excluding* details when trying to maximize acquisition of important points:

(1) There are tens of thousands of facts

in a text; we cannot expect students to memorize all of these facts and they don't.

(2) Students will have to time share, that is, devote some of their processing time to unimportant facts.

(3) It is harder for readers to appreciate or extract the important points if they are embedded in details.

(4) Hierarchical analyses of text-structure assert that details are subordinate in the representation to the main points and therefore cannot help the student remember the main points. Access to details is through the higher level nodes and thus dependent on recovering the main points.

The arguments for *including* details in an attempt to maximize acquisition of important points can be summarized as follows:

(0) Details may provide needed arguments to convince the reader of the verity of an assertion. However, this and other functions of illustration are not relevant to the discussion of how to best learn specific facts. Thus we have called this reason 0.

(1) Embellishments may allow the reconstruction of the main points. The details imply the main points although the converse is not true.

(2) Sometimes access to central points may be available only via some details. We were inclined to believe this second point of view given the importance we attribute to elaboration in other contexts (Anderson, 1976; Anderson & Reder, 1979; Reder, 1976, 1979). In an informal poll of colleagues we also found that the much more frequent prediction was that text would prove to be better than summary—at least at delayed testing.

EXPERIMENTS 1 AND 2

To discover whether inclusion of details is advantageous or not, we decided to perform a pair of experiments that would test subjects' memory for information acquired via text and via summaries immediately after study and at a delay of 1 week. The two experiments differ only in the amount of study time, so they will be described to-

gether. Two standard introductory college texts were selected as the basic material, one *An Introduction to Descriptive Linguistics* by Gleason (1967), the other *The Geography of Modern Africa* by Hance (1975). Neither text requires prior knowledge on the part of the student in that content area. In both cases, Chapter 1 was the material students studied. Summaries were written for both chapters. These summaries were roughly one-fifth the length of the originals and were about 1000 words long. The questions we chose to ask the subjects could all be answered on the basis of the summaries. The questions were true-false and half of the trues and half of the falses could be answered by retrieving a simple assertion provided in the summary. The other half required that the reader combine statements presented in the summary. The former type are called *direct* questions; the latter, *indirect* questions.

Method

Materials. The text material had been photocopied from the original textbooks. The summaries were written to restate the main points in as compact a fashion as possible. We could not find a published procedure for generating the summaries that was sufficiently general in application and clearly specified to allow one to represent complex texts such as these. All systems for representing text seem to require involved, tedious techniques that are difficult to use for one paragraph. Trying to adapt them to chapters over 13 pages long would border on the impossible.

Rather than try to adapt a procedure used to generate representations of paragraphs to one that generates summaries of chapters, we used our intuitions to rewrite the material in a concise fashion. Several people were involved in the exercise of generating summaries. Before a summary was completed, a consensus had to be reached such that all people involved (at least two) thought that the summary reflected the important points in the passage. We also attempted to analyze the nature of

our implicit rules used to generate the summaries. These rules look like a composite of van Dijk's (1977) summarization rules for shortening sentences and a rule for deleting "lower level" paragraphs and propositions where lower level is defined by a Meyer (1977) type of text analysis. Below is a portion of the original Gleason text, the corresponding portion of the summary and a few of the questions used to test the information.

1.1 As you listen to an unfamiliar language you get the impression of a torrent of disorganized noises carrying no sense whatever. To the native speaker it is quite otherwise. He pays little attention to the sounds, but concerns himself instead with some situation which lies behind the act of speech and is, for him, somehow reflected in it. Both you and he have failed to grasp the nature of the phenomenon. Neither the casual observer nor the usual native speaker can give any real information about a language. To be sure, some people, Americans perhaps more than most others, have decided notions about language. But the ideas held and discussed come far short of giving a complete picture of the language and sometimes have very little relationship to the facts. Even people with considerable education are often wholly unable to answer certain quite simple questions about their language. For most people language is primarily a tool to be used, rather than a subject for close and critical attention.

It is probably well that it is so. Yet there are important human problems into which language enters intimately and on which it exerts such a profound influence that an understanding of its mechanisms would contribute materially to their solutions. Moreover, every phase of human activity is worthy of study. Thus, for practical reasons, as well as to satisfy man's innate curiosity, language deserves careful and intelligent study. Gleason (1967, p. 1)

Corresponding Summary:

- An unfamiliar language sounds like a meaningless torrent of noises. (1)
- A native speaker attends to the meaning, not the sound of what he hears. (2)
- The layman, untrained in linguistics, does not understand the true nature of language. (3)
- The layman uses language as a tool and does not study it analytically. (4)
- Language is worth studying because an understanding of language can help in

the solution of important human problems. (5)

Questions:

Direct True: A nonlinguist does not appreciate the true nature of language.

(From 3)

Indirect False: A native speaker attends to the phonemes of his language.

(This sentence can be rejected using sentence 2 and subsequent information that phonemes are the basic sounds of language.)

The summary statements were typed on plain sheets of paper, keeping the section headings of the original text. No paragraph structure or interstitial material was maintained, however. Each sentence started on a new line. The African Economic Geography text was approximately 4300 words and its summary 1000 words. The linguistic text was about 4900 words and its summary 800 words.

For each topic, 32 questions were written, half interrogating specific assertions in the material (direct questions) and half requiring integration of several points (indirect questions). To facilitate collection of reaction times, the items required only simple true-false decisions. Half of the direct and half of the indirect questions were false. The items were selected such that a group of Yale undergraduates who had not read the material answered the questions with chance accuracy, viz., scored at about 50%. The test items were never verbatim repetitions of statements in the summary or text. Nor could the true questions be considered close paraphrases of the studied statements. A direct true does clearly follow from an assertion made in a summary, but unless the reader has bothered to comprehend the material, he or she is unlikely to do better than chance at recognizing the questions.

Procedure. Each subject studied both the linguistics material and the African material, one in the original text form and the other in the summary form. Subjects in Experiment 1 studied the text for 20 minutes and the summary for 20 minutes. Some

subjects complained that 20 minutes was not sufficient time to read the text; so in Experiment 2, subjects were given 30 minutes to study each type of material. Whether subjects studied the text material first or the summary material first was randomly determined as was the assignment of the topics to the two study conditions.

Subjects were given the reading matter to study in a small, private cubicle that contained a video terminal connected to a PDP 11/40 computer. The terminal indicated to the subjects how much study time was left.

After studying each topic, subjects were tested with a series of 16 statements. These 16 statements were composed of four indirect true, four indirect false, four direct true, and four direct false as defined earlier. Thirty-two questions were constructed, but only half of each type were (randomly) selected to be tested immediately after studying. The other half were tested 1 week later. Reaction times were surreptitiously recorded from the onset of the statement until the subject responded as to whether the assertion was true or not as defined by the material studied. In calculating mean reaction times, latencies greater than 20 seconds were truncated to 20 seconds. This was done to avoid the distorted means that might arise if a subject sat and thought for minutes about a question.

After answering questions about the first topic studied, subjects studied the second

topic. When subjects were brought back a week later to answer the remaining 16 questions on each topic, they were tested on the two sets in the same order as they had been studied. Order of statements in both the initial testing and the delayed test were randomized.

Subjects. Thirty-nine subjects were tested in Experiment 1 and twenty-four in Experiment 2. They were recruited from the undergraduate population at Yale and served either to earn credit in the introductory psychology course or to earn \$5.00. The experiments were run in two sessions. The first lasted an hour or more. The second, a return test 1 week later, lasted about 10 minutes.

Results

Table 1 presents the results from the two experiments. There is a slight advantage for subjects who studied the material for 30 minutes, but the pattern is essentially the same. In fact, subjects are only 1.8% better in Experiment 2 for the 10 extra minutes of study time. Therefore, we will present the data of all 63 subjects averaged together. Table 2 presents the data collapsed over Experiments 1 and 2, but partitioned into immediate and delay conditions. Each cell of the matrix represents 252 observations.

TABLE 1
PROPORTION CORRECT ON QUESTIONS FROM
EXPERIMENTS 1 AND 2 AS A FUNCTION OF TYPE
OF STUDY AND TYPE OF QUESTION

	Experiment 1 (20-minute study time)	
	Summary	Text
Direct	.772	.625
Indirect	.720	.659
	Experiment 2 (30-minute study time)	
	Summary	Text
Direct	.789	.638
Indirect	.732	.691

TABLE 2
ACCURACY AND LATENCY PERFORMANCE ON
QUESTIONS FROM EXPERIMENTS 1 AND 2 AS A
FUNCTION OF TYPE OF QUESTION, TYPE OF
STUDY, AND DELAY

	Proportion correct			
	Immediate		Delay	
	Summary	Text	Summary	Text
Direct	.839	.651	.718	.607
Indirect	.752	.707	.700	.647
	Reaction time (seconds)			
	Immediate		Delay	
	Summary	Text	Summary	Text
Direct	10.64	11.64	10.84	11.13
Indirect	12.86	12.80	11.84	11.86

An analysis of variance was performed on the data using as factors subjects, form of study (text vs summary), truth (true vs false), question type (direct vs inference), and delay. The standard error of the means in Table 2, based on the overall subject-by-condition interaction, is 2.9%.

There are significant effects of study [$F(1,38) = 15.13, p < .001$] and delay [$F(1,38) = 16.53, p < .001$]. The advantage of summary over prose does not seem to diminish with delay. The overall effect of truth is not significant and we collapsed over it since differences between true and false may only reflect response biases.

There is a sizable interaction between the dimensions of study and statement type [$F(1,38) = 8.00, p < .01$] such that subjects answer direct questions better than indirect questions when the material is studied in summary form, but answer indirect questions better when the material is learned in the original text form. However it is still the case that performance on indirect questions is better for material learned in summary form than in prose form, i.e., there is a main effect of type of study over and above the interaction.

Table 2 also presents the judgment times (in seconds). These data were analyzed using the same factors as the percentage correct data. The same pattern of results is obtained, although the differences are not all reliable. Subjects were significantly faster in responding at a delay [$F(1,38) = 5.38, p < .05$]. Subjects were also faster to answer direct questions than indirect [$F(1,38) = 58.77, p < .001$]. Subjects seem to be faster somewhat to answering questions about material studied in summary form, but this result is not reliable.

Judgment times were collected surreptitiously, as we were principally interested in accuracy. Unlike studies where reaction time is the primary dependent measure, it would be senseless to report here only RTs for correct responses. Error rates in this study are quite high as we expected and clearly a great deal of thought goes into

making those judgments which turn out to be incorrect. We are interested in the differences in times to arrive at conclusions as a function of study condition regardless of whether the knowledge is adequate to arrive at the correct conclusion.

We also analyzed the results as a function of type of material learned, viz. whether the data looked the same for the linguistics chapter and the African economic geography chapter. Subjects performed slightly better on the linguistics chapter than on the African chapter, however the same pattern of results was obtained for both chapters.

EXPERIMENT 3

Experiments 1 and 2 equated the summary and text material on study time in order to assess which format was best for communicating main points. The conclusion seemed to be a clear victory for the summaries with an overall accuracy of 75.2% as compared with an accuracy of 65.3% for material studied in text form. However, the question of retention seems to be at least as theoretically interesting as initial acquisition and it occurred to us that prose might show a better retention function.

In order to discover whether material acquired from original texts would be more resistant to forgetting, we needed to equate initial performance. With this in mind, we ran a third experiment giving subjects only 15 minutes to study the summaries and 45 minutes to study the text. In other respects, the procedures and materials were identical with the first two experiments. The experiment used 31 subjects recruited from the same population as before and they could receive either credit or \$5.00 payment.

Results

Table 3 provides the percentage recall data from the experiment as a function of form of study, type of question and delay. The difference between summary and text is only 2.5% in immediate test as compared with 11.7% in Experiments 1 and 2. These data were analyzed in the same way, using

TABLE 3
ACCURACY ON QUESTIONS FROM EXPERIMENT 3
AS A FUNCTION OF TYPE OF STUDY, TYPE OF
QUESTION, AND DELAY

	Proportion correct			
	Immediate		Delay	
	Summary	Text	Summary	Text
Direct	.76	.68	.68	.65
Indirect	.73	.76	.67	.59

the same factors as the data from Experiments 1 and 2. The standard error of the means is 4.2%.

There is no longer a significant difference between summary and text in the immediate condition, although there is still a slight advantage for material studied in summary form. Now one can ask whether the text material is better retained. There is no significant interaction of delay with type of study [although there is a significant effect of delay— $F(1,30) = 20.22, p < .001$]. To the extent that there is an advantage of retention for one study form, the evidence is in favor of summaries again.

In Experiments 1 and 2, it was noted that there was an interaction between type of question asked and study form. In those experiments, the interaction seemed to diminish slightly with delay, but there was not a significant three-way interaction with delay. In the current experiment, the interaction of question type and study form is present only in the immediate condition. At a delay, indirect questions are not answered as well as direct questions in both study conditions.

Discussion

So far, the data have not indicated any circumstance where subjects perform better having studied the original text than having studied a summary of it. There was scant evidence that indirect questions are answered better than direct questions when material is studied in text form. However, this effect was weak statistically and not so large absolutely to cause a reversal with the

summary condition. There may be a temptation to attribute the poor performance in the text condition to the particular materials chosen. For instance, perhaps the texts were poorly written. This seems unlikely as these texts have been adopted in a number of college courses. There is also at least one published study (Newsom & Gaité, 1971) that has also found better memory for summary material (in a somewhat different paradigm).

Nonetheless, to ensure that these results could not be accounted for totally by inherent properties of the materials used, an experiment was performed that removed the memory component from the task. This would indicate to what extent the results could be attributed to the ease of question answering from the two sets of materials. Later experiments would explore further the generalizability of the results to other materials.

EXPERIMENT 4

Subjects were given the same questions to answer in this experiment as other subjects were given in previous experiments. They were also given the same materials, viz. one chapter in the original text form and one in the summary form. The principal difference was that the material they were to study was before them while answering the questions and they were encouraged to go back and look up any answer about which they were uncertain. The goal of this experiment was to eliminate the memory component of the task, and discover how much of the difference between the two study conditions was attributable to ease of discerning the relevant information from the study material.

It should be pointed out that this procedure does not really eliminate the role of memory in the task. Subjects were told to read the material first before answering the questions. They did this, but tended not to look back at the passage to answer the questions. After a few pilot subjects were run, a decision was made to pay subjects

for correct responses and penalize them for incorrect answers in order to motivate them to look back through the material. Even with a monetary incentive, subjects did little hunting. In any case, the results from this experiment give some idea of the extent to which the results can be attributed to comprehension rather than memory failure.

Subjects. Twenty subjects were recruited from the Yale undergraduate population for an experiment that lasted under 2 hours. They were paid \$2.50 plus a bonus of \$.05 for each question they answered correctly, minus \$.05 for each question answered incorrectly.

Procedure. The same materials were used as in Experiments 1, 2, and 3. However, subjects were given unlimited time to read the study material. After reading through the material to their satisfaction, they were given the 32 statements for that topic on a sheet of paper and were encouraged to go back and check their judgments about the truth of these statements. It was emphasized that it was important to be certain about their judgments and to find the supporting information when necessary. Subjects did search through the material at least some of the time. Each subject read both topics, one in text form and one in summary form. All possible orders of topics and conditions (four possible) were used equally often.

Results and Discussion

Table 4 presents mean accuracy for direct and indirect questions on material read in summary form and in text form. Subjects did not do all that well even with the text in front of them. This caused us to doubt whether our designated "correct" answers were always correct. In an informal study,

with a few subjects, we pointed to the portions of the text from which the questions were taken. We obtained over 95% agreement with our subjects in this case. The subjects' problem in this study was that they did not know which portion of the text was relevant to the question. Thus, it is clear that in all conditions subjects are making errors primarily because they are not always searching when and where they should (cf. Hayes-Roth & Walker, 1979).

It is unclear what portion of the subjects failure in the memory conditions is due to similar factors. Perhaps they had the information committed to memory for answering the question but failed to recall it in response to the question. This seems particularly plausible with respect to the indirect questions.

In any case, we took the conservative option of assuming that every time the subject failed in this experiment he would have failed in a memory experiment. We corrected the performance from the previous experiments with the data from this experiment under this conservative assumption. To do this we assumed that failure to recall in the previous experiments was a mixture of failure to notice the *relevance* of information in memory which could be directly estimated from this experiment and "true memory failure." Let p_n be the probability of failure to notice relevance of information. Let p_o be the *corrected*, observed probability of failure to recall in Experiments 1 and 2 and p_m be the probability of a memory failure. We obtained estimates of P_o from Experiments 1 and 2 and of P_n from Experiment 4 by doubling the observed error rates. The new error rates derive from the standard guessing model which assumes that half of the failures result in errors and the other half result in lucky guesses. The relation between these corrected probabilities is

$$p_o = p_n + (1 - p_n)p_m$$

and so p_m can be estimated

$$p_m = \frac{p_o - p_n}{1 - p_n}$$

TABLE 4
ACCURACY ON QUESTIONS WITH "NO
MEMORY REQUIREMENT"

	Summary	Text
Direct	.85	.78
Indirect	.80	.72

Table 5 gives the estimated probabilities of a memory success ($1 - p_m$) for Experiments 1 and 2.

The pattern obtained by analyzing the data correcting for failure to notice the relevant information tells much the same story as before. Subjects can answer questions better if they have the summary material before them. The estimate of differential performance without memory constraints undoubtedly contains effects due to memory. Subjects often did not look back to the materials when answering questions. Therefore correcting memory data by dividing by the results from this control experiment will attenuate a real difference. Nonetheless, there still remained an advantage for the summary condition.

EXPERIMENT 5

The data do not support the notion that the reason details are included in a text is to enable readers to retain the central points of a passage better. Conceivably a more subtle benefit might accrue with the inclusion of details. Details provide the reader with a richer, more elaborate structure of the knowledge and perhaps this elaborated structure helps that person better acquire *subsequent information*. This hypothesis can be tested by looking at the extent to which subjects perform better on learning a set of facts when other related facts, learned previously, were acquired in text form rather than summary form.

New materials were used in this experiment for several reasons. First, the experiment required text book chapters that could be easily split in half so that one-half could

be learned in one form (text vs summary) and the other half in the other form. Second, this provided a test of the replicability of previous results with new materials.

Method

Subjects. Thirty-two subjects from the Yale University community were recruited during the summer session. They were paid \$7.50 for their participation.

Materials. Again, introductory textbooks were used. One chapter dealt with the money supply taken from an introductory economics text called *Macro Economics* by Lloyd Reynolds (1976), the other chapter, on the Russian Revolution, is from a history text called *Russia: A Short History* by M. P. Florinsky (1969). Both chapters have a natural stopping point in the middle which served to divide them in half. The first and second halves of the economics were about 4000 and 4500 words, respectively. The first and second halves of the history chapter were 5000 and 6000 words. Subjects were asked questions about each half chapter after studying that material.

Again, summaries were written to restate the main points in as compact a fashion as possible. Each summary (for a half chapter) was about 900 words. The section headings of the original text were retained, but there was no paragraph structure or interstitial material for the summaries. Only the critical sentences were typed on plain paper. Subjects received photocopies of the original chapters in the text condition. True-false questions were constructed with the constraint that they could be answered on the basis of the summaries alone. Since the important findings did not depend on the distinction between direct and indirect questions, the distinction was dropped for these topics. All questions were pre-tested and selected only if Yale undergraduates who had not read the material performed at chance (50% accuracy) on them. There were 20 questions, 10 true and 10 false for each half topic.

Procedure. Each subject studied both the economics and the history material. The

TABLE 5
ESTIMATED PROBABILITIES OF MEMORY SUCCESS
USING CORRECTED MEMORY SCORES DIVIDED
BY CORRECTED "ABSOLUTE" SCORE

	Immediate	Delay
Summary	.91	.64
Text	.72	.51

second half of a topic was always studied after the first half. There were 32 different conditions in this experiment made by all possible combinations of summary or prose in the four half topics and the two possible orders, economics or Russian studied first.

Each half topic was studied for 30 minutes. Subjects were asked 20 questions immediately after reading each half.

Results

The data from this experiment are displayed in Table 6 averaged over order, truth, and type of text. Separate statistical analyses were done for the first text studied and the second.² We will report the results of these analyses combined. Focus first on the top portion of the table. This lists the overall performance for the combined halves of both topics studied in summary form and studied in text form. Once again, there is an advantage for the summary condition although the effect seems a little weaker. The effect is nonetheless significant [$t(96) = 2.80$]. The fact that the effect is not as large may be due to a difference in the materials selected.

² Because of the nature of our design, we had to do separate analyses for the two topics read by a given subject. The factors were topic, form of study first half, form of study second half—all between. True vs False was a within-subject factor.

TABLE 6
PROPORTION CORRECT IN EXPERIMENT 5:
TRANSFER EXPERIMENT

Performance averaged over both halves			
	Summary	Text	
	.758	.704	
Performance on second half as a function of study form on first half			
Current form	Previous form		
	Summary	Text	Average
Summary	.782	.721	.752
Text	.714	.737	.726
Mean	.748	.729	.739

The bottom portion of the table presents the results of major interest. It displays the proportion correct on the second half of each topic as a function of how the second half was studied and as a function of how the first half was studied. There is a marginally significant interaction [$t(96) = 1.61$] such that a person is better off studying the second half in the same form as the first half was studied. Over and above this, however, there is no advantage for having studied the first half in prose form as we had anticipated. If anything, there is an advantage here too for having learned material in summary form [but clearly not significant— $t(96) = .92$].

EXPERIMENT 6

Some of the advantage for the summary conditions might be due to the fact that subjects' perceptions of the critical points differ from the writers of the summaries and therefore in the text condition they are not focusing on the facts to be tested. Kintsch (personal communication) has found in the past that students' note-taking of his lectures reveals a different organization of ideas such that different points are considered the central ones, (see also Kintsch & Bates, 1977).

In order to test the notion that the advantage of the summary condition was due to subjects inability to isolate what we thought the important points in the text, we replicated the last experiment and underlined the main points in the text. In all other respects, the two experiments were identical. All the points in the summaries could be found in the text and these main points were underlined on the photocopies of the chapters. Thirty subjects were taken from the same population.

Results

The results from this experiment are presented in Table 7. The top portion of the table gives the means for the two critical conditions collapsed over materials and halves of chapters. The effect of type of study is attenuated [$t(90) = 1.69$], but the

TABLE 7
PROPORTION CORRECT IN EXPERIMENT 6:
TRANSFER EXPERIMENT WITH THE
MAIN POINTS UNDERLINED

Performance over both halves			
Summary	Text		
.734	.718		
Performance on second half as a function of study form on first half			
Current form	Previous form studied		
	Summary	Text	Mean
Summary	.767	.687	.727
Text	.758	.717	.738
Mean	.763	.702	

advantage for summaries remains. The fact that the effect is smaller should not be surprising in that some subjects may have simply read the underlined portions of the text material, i.e., treated the text condition as if it were the summary condition.

The center portion of the table presents the critical analysis. Again there is an interaction such that subjects perform best learning the second half in the same mode that they learned the first half [$t(90) = 1.53$]. Combined, these two experiments yield a significant interaction between first and second form [$t(186) = 2.22$]. There was also a marginally significant advantage in this experiment of having studied the earlier text in summary form [$t(90) = 1.75$]. In combination, the two experiments also display a marginally significant effect in that direction [$t(186) = 1.89, p < .1$; two-tailed].

EXPERIMENT 7

Having been unable to find an advantage for prose in any circumstance, we made one last attempt to see if perhaps the long-range retention of prose might not be superior to acquisition from summaries. We brought back as many subjects as we could from Experiments 1, 2, and 3. The original three experiments had been performed from January through May of 1977 while the retention tests were administered from

November 1977 through May 1978. Thus, the delays varied from 6 to 12 months. The variation in delay is not serious as each subject serves as his own control.

Method

Subjects. It was not possible to achieve even a 50% return rate. Some subjects had left Yale, others were impossible to contact, and still others refused to return. Of the 39 subjects in Experiment 1, there were 16 returns; of the 24 in Experiment 2, there were 16 returns, and of the 31 in Experiment 3, there were 18 returns.

Materials. New true-false questions were created for both topics used in the experiments. Unlike the original set, no distinction was made between direct and indirect questions. But like the original questions, they were tested on pilot subjects who had not read the materials to assure that they would have only 50% (chance) accuracy of a correct response.

Procedure. Subjects were not told prior to their return why we needed them for our experiment. That is, we did not tell them that they would be asked more questions on the material they studied. In this way, we hoped to minimize any rehearsal prior to the retention test.

As in the original three experiments, the questions were presented to the subject via computer-driven CRT screen. They were tested on the two sets of material in the same order as originally studied. It was carefully explained to the subjects that they were still to try to judge the truth of these statements with respect to the material that they had originally studied. While many subjects claimed to remember little or nothing from the original study episode, it was at least the case that they were able to remember which material was studied in summary form and which in text form.

Results

The results from this experiment are presented in Table 8 along with the comparable data from Experiments 1, 2, and 3 using the data of only those subjects we

TABLE 8
PROPORTION CORRECT FOR SUBJECTS IN
EXPERIMENTS 1, 2, 3 WHO PARTICIPATED
IN EXPERIMENT 9

	Summary	Text
Immediate	.793	.693
1 week	.675	.600
6 months/1 year	.595	.575

were able to bring back. The data are presented for summary and text conditions as a function of the three levels of delay, viz. immediate test, test at 1 week, and test at 6–12 month delay.

Even at a delay of up to 1 year, there is still some advantage for the summary condition. There is an interaction between study form and delay [$F(4, 188) = 2.68, p < .05$], and the difference between the text condition and the summary condition is not significant at the longest delay. These effects may be due to the fact that performance is approaching the “floor” which is 50%. That is, the interaction and the lack of a significant difference might be due to the fact that true differences cannot be manifest since performance using percentage correct cannot go any lower.

GENERAL DISCUSSION

The data from seven experiments have been presented all of which argue that learning material from summaries is at least as good as reading the original text. People’s ability to recognize important facts about a topic after studying it, regardless of the delay between study and test, is superior when the information is learned from a summary. New information is learned better (measured by one’s ability to answer questions) if information learned earlier on a related topic was learned by reading a summary.

Our initial expectation was that the embellishments would improve retention since they provide a redundant coherent structure. Apparently the total time law and the notion of having to time share between main points and details is a more accurate

way of understanding the situation. It is possible that subjects are providing their own embellishments while studying the summaries, but that is pure speculation. The fact that texts do not have a retention advantage over summaries is something of a consternation for our elaboration analysis of memory. Of course, because of the uncontrolled structure of textbooks, the reasoning is loose that goes from the elaboration hypothesis to a prediction of an advantage for textbooks. We plan to follow up these results with experimentation on more carefully controlled material.

In the introduction we briefly reviewed the positions of some prominent theorists with respect to the role of details in supporting memory for the central points. We concluded that the dominant view of “text grammarians” is consistent with the notion that details do not support memory for the central, important ideas. In hierarchical representations of text, details can be retrieved only by first retrieving the higher level points. The data are consistent with this position.

There are undoubtedly other desired outcomes besides fact learning that are intended from studying a text. Conceivably, a text serves these other purposes better than would a summary. We have not shown that students better appreciate the nature of a field of study by reading a summary. It is also quite likely that outside of a laboratory setting students would read the summary only once. In that case, it may be to their advantage to read the prose instead.

Perhaps part of the reason that subjects learn material better from summaries is due to an unwillingness or an inability to isolate the main points and only skim details that are not of great interest. Research has indicated that better readers are more prone to highlight passages and that when all students are required to highlight, only good readers highlight the important points. Poor readers tend to highlight points at random (Smiley, Oakley, Worthen, Campione, & Brown, 1977).

It had occurred to us that some subjects

might benefit more from reading summaries than other subjects. In fact, there is some evidence (Pask & Scott, 1972) that different groups of people learn the same set of material better when it is structured in a manner consistent with their respective competencies. It seemed reasonable to suppose that poor readers might benefit more from reading summaries than good readers who can discern the important information from the text. To test this notion, the data from Experiments 1 and 2 were divided in half such that the subjects with the highest accuracy were in one group and subjects with the worst accuracy in the other. We then reanalyzed the data for each group to see if the pattern of summary versus text study conditions would be different for the good and bad subjects. We found no difference in pattern. So it seems that the findings of better performance when reading summaries generalizes to students of varying abilities.

Some of the blame for failing to extract the important points could be attributed to bad writing. The author may not always make clear what the central ideas in the message are.

Another reason we might have obtained these results pertains to the nature of the dependent measure employed. Conceivably a recognition test is not a sensitive enough measure to demonstrate the advantages of a rich, elaborated structure like the text form of the study material. We chose not to use free recall as a measure because we felt that the prose condition would be at a disadvantage. In scoring the recall protocols, only that information contained in the summaries could be counted. Output interference would put in text condition at a large disadvantage. That is, subjects would recall a lot of information that would not count in scoring but probably would prevent them from retrieving other information that would count. Nevertheless, research is currently underway where the main points in the text are underlined and subjects are told that they need only recall the underlined portions.

REFERENCES

- ANDERSON, J. R. *Language, memory, and thought*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.
- ANDERSON, J. R., & REDER, L. M. An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.
- BARTLETT, B. J. *Top-level structure as an organizational strategy for recall of classroom text*. Unpublished doctoral dissertation, Arizona State University, 1978.
- CROTHERS, E. Memory structure and the recall of discourse. In J. Carroll & R. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston, 1972.
- FLORINSKY, M. P. *Russia: A short history*. London: Macmillan, 1969. Pp. 411-437.
- FREDERICKSEN, C. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975, 7, 371-458.
- GLEASON, H. A. *An introduction to descriptive linguistics*. New York: Holt, Rinehart & Winston, 1967. Pp. 1-13.
- HANCE, W. A. *The geography of modern Africa*. New York: Columbia Univ. Press, 1975. Pp. 5-20.
- HAYES-ROTH, B., & WALKER, C. Configural effects in human memory. *Cognitive Science*, 1979, 3, 91-118.
- KINTSCH, W. *The representation of meaning in memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
- KINTSCH, W., & BATES, E. Recognition memory for statements from a classroom lecture. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 150-159.
- KINTSCH, W., & VAN DIJK, T. A. Comment on se rappelle et on resume des histoires. *Langages*, 1975, 40, 98-116.
- MANDLER, J. M., & JOHNSON, N. S. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, 9, 111-151.
- MEYER, B. J. F. *The organization of prose and its effect on recall*. Amsterdam: North-Holland Publishers, 1975.
- MEYER, B. J. F., & McCONKIE, G. W. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, 65, 109-117.
- NEWSOM, R. S., & GAITE, A. J. Prose learning: Effects of pretesting and reduction of passage length. *Psychological Reports*, 1971, 28, 123-129.
- PASK, G., & SCOTT, B. C. Learning strategies and individual competence. *International Journal of Man-Machine Studies*, 1972, 4, 217-253.
- REDER, L. M. *The role of elaborations in the process-*

- ing of prose*. Unpublished doctoral dissertation. University of Michigan, 1976.
- REDER, L. M. The role of elaborations in memory for prose. *Cognitive Psychology*, 1979, 11, 221-234.
- REYNOLDS, L. G. *Macroeconomics: Analysis and policy*. Homewood, Illinois: Irwin, Inc., 1976. Pp. 191-210.
- RUMELHART, D. E. Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.
- SMILEY, S. S., OAKLEY, D. D., WORTHEN, D., CAMPIONE, J. C., & BROWN, A. L. Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. *Educational Psychology*, 1977, 69, 381-387.
- STEIN, N. L., & GLENN, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *Discourse processing: New directions*. Norwood, N.J.: Ablex, in press.
- THORNDYKE, P. W. *Cognitive structure in human story comprehension and memory*. Unpublished doctoral dissertation, Stanford University, 1975.
- THORNDYKE, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 1977, 77-110.
- VAN DIJK, T. A. Semantic macro-structure and knowledge frame in discourse comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

(Received August 13, 1979)