THE EFFECT OF FORMAT AND STRUCTURE OF TEXT MATERIAL ON RECALLABILITY

CARL MARTIN ALLWOOD, TORBJORN WIKSTROM and
LYNNE REDER *

Prior research (Reder and Anderson 1980, 1982) has shown that acquisition of new information can be facilitated by a format that *summarizes* the important points as compared with a more traditional format, such as a textbook chapter. The previous results used recognition (true/false tests) and cued recall (who did what?) measures. The experiment reported here investigated whether the detrimental effects of details in texts are compensated by their facilitory effects in providing additional retrieval routes. This was tested by using a free recall paradigm. Once again, summaries resulted in superior memory for the main points. Possible explanations and pedagogical implications are discussed. The quality of elaborations in texts and the ease of generating well organized representations from different formats may both affect performance.

With the seemingly exponential growth in the amount of written material in almost every discipline of interest, the question of how to utilize a student's time most effectively to absorb this information has become quite important. The research to be reported here concerns an investigation of the optimal format for presentation of factual or textual information. Previous studies by Reder and Anderson (Reder in press a; Reder and Anderson 1980) have found again and again that subjects are better able to answer questions if they have learned the material from reading a summary than if they had read the original text material. The effects were manifest for both tests of direct knowledge and those that involved inferential reasoning. Summaries were also found to facilitate (when compared with texts) subsequent learning of related material, regardless of the new material's format. In other words, people learned a chapter better if they had just studied a related chapter in summary form than in text form. The pattern of summary superiority emerged when true/false questions were employed and when cued recall was the dependent measure. These effects were found immediately after study, one week later, and from six months to one year later.

Other studies of Reder and Anderson (1982) have also shown an advantage

for materials that omit embellishment when compared with the same central ideas presented with the elaborative, supportive information. In those studies, presentation rate and format were constant across conditions. Several explanations are possible for why elaborations hurt retention: one is that elaborations compete with the main points in terms of information processing time and thereby reduce the probability that the central points will be stored. One experiment controlled the study time of the central points so that all study time for embellishments was additional. The effects were as large as in other studies, suggesting that the advantage of summaries cannot be explained by the total time law (Bugelski 1962; Cooper and Pantle 1967).

A second possibility is that embellishments interfere with the main points at time of retrieval. Conceivably the additional facts make the representational structure in memory more cluttered. A more cluttered structure might make it more difficult to read or interpret the central facts in memory. That is, it is difficult to detach the central points from the related details to which they are attached. However, these redundant connections in memory might prove facilitory in a recall task. In recall, additional paths to the *relevant* conceptual structure would be useful. In a recognition or cued recall task, access to the appropriate concept nodes in memory is guaranteed. So the following trade-off might occur: additional facts attached to the relevant concept facilitates access to it; additional facts attached to the concept interfere with retrieving the appropriate one (e.g., Postman 1961; Anderson 1974). In recognition or cued recall, the benefit of extra pathways to the relevant concept would be minimal. In a free recall task, however, the interfering effects of extra facts (embellishments) might be overcome by the advantages of extra retrieval routes to that topic node. The present study was an attempt to see whether standard, embellished texts would produce better performance than their summaries in a free recall situation, or at least whether the advantages of summaries would be attenuated in this task.

A central problem in a task such as this is determining an appropriate method to compare performance in the two conditions. In a true/false task or a cued recall task, the experimenters can control the questions so that they tap information studied in the summaries as well as in the text. The issue had always been how to maximally learn the important ideas (i.e., those main ideas contained in the summaries) and therefore the embellishments omitted from summaries were not tested. However, in a free recall task, the experimenters cannot so easily constrain subjects who have read the whole text to recall only the summary material, to the exclusion of the details. It would not be fair to "count" the recall of any fact that had been read since subjects in the text condition would have five times as many facts that they could recall. However, not counting recalled facts that were not deemed important would put subjects in the text condition at an enormous disadvantage. First, they may not appreciate what the experimenter perceives as the important ideas and there-

fore recall other facts; second, they may suffer from an output interference problem, such that they might know a specific important fact but they forget to recall it after recalling many details. The notion of output interference in recall is a robust phenomenon (Smith 1971).

In addition to the output interference problem, there can be a bookkeeping problem such that subjects may worry what type of facts they have recalled or are recalling and whether they have already recalled that fact before. The issue of whether a particular fact has already been recalled will be common to both the summary and the text subjects. In an attempt to minimize these problems we adopted the following solutions: first, we highlighted the facts in the original text that we considered to be the main points. These were essentially the same statements that were presented to the subjects in the summary condition and the facts we wanted to "count" or score in the recall protocols. In previous research of Reder and Anderson, underlining the main points in the original text attenuated the advantages for the summary condition but did not eliminate them. We assumed at the time that to the extent that underlining improved performance in the text condition, subjects were occasionally adopting the strategy of only reading the underlines, i.e., treating that condition as the summary condition. Therefore, we expected to aid subjects in the text condition by providing the highlighting or underlining of the critical points. Less processing had to be devoted to determining the important points.

Second, at the time of free recall, subjects who had read the text were instructed to try to recall all of the underlined points. They were clearly told that they were allowed to recall any material although we were not interested in that material. We did not want subjects to be inhibited from recalling a fact because they feared it had not been underlined nor to devote much attention to trying to discriminate at test underlined statements from non-underlined statements. To the extent that subjects do not recall non-underlined materials, they should not suffer from more output interference than the summary subjects.

Third, we took two measures of recall performance. One measure compared overall number of ideas recalled that had been underlined in the text with the number of ideas that had been recalled from the summary. We also compared total number of unique ideas recalled from the text, regardless of underlining, with recall from summary. This was done in case our attempts to minimize output interference were unsuccessful.

Method

Materials

Two of the passages used by Reder and Anderson (1980) were employed. These were the first chapters of two textbooks, *An Introduction to Descriptive*

Linguistics by Gleason (1967: 1–13) and *The Geography of Modern Africa* by Hance (1975: 5–20). Both were presented in their original forms and in their summary forms constructed by Reder and Anderson (1980). Summaries restated the main points and were approximately one-fifth the length of the original. The prose version of the African text was approximately 4,300 words long and its summary approximately 1,000 words long. The corresponding figures for the linguistics material were 4,900 and 800 words. The parts in the original chapters corresponding to the information in the summary versions were underlined.

Design and procedure

The principal within subject factor in this design was format of presentation of text material, with two levels, original underlined text and summary. The two topics were studied by all subjects. Order of presentation of topic and assignment of format type to topic were counterbalanced by means of a Latin-square design.

The study materials were given to the subjects in envelopes. Before reading the texts, the subjects were informed that they would be given a memory test on the topics studied. They were further informed that they need not remember the exact wording, just the ideas. Subjects were instructed when to open the first envelope and begin to study, when to stop studying, when to begin studying the second packet of material, etc. They were given one half-hour for each topic.

After studying both topics for a total of one hour, subjects were instructed to recall as much as they could from the first passage for one half-hour. At the end of this half-hour recall, subjects were then asked to try to recall the second passage for the next half-hour.

Subjects who were to recall the summaries were asked to write down as much as possible of all statements from the summary, and the subjects who were to recall the text version were asked to "try to remember as much as possible of the things underlined in the text. If you are not sure whether an idea you remember was underlined or not, then write it down anyway". This last instruction was included in order to minimize the bookkeeping problem discussed above.

Subjects

Sixteen Carnegie-Mellon University undergraduates participated to fulfill a course requirement. The experiment took approximately 2 hours and subjects were given 2 credits (out of a required 3).

Results

The primary dependent measure was proportion of central ideas recalled. Before describing the results, it is essential to indicate the manner in which recall was scored. The underlined facts in the text version of each topic correspond quite closely to the statements in the summary version, and therefore we will describe how the scoring worked for the summaries. Each statement in a summary was assigned a number of points corresponding to the number of independent ideas it seemed to contain. Each complete proposition received 1 point and a subordinate proposition received $\frac{1}{2}$ (one-half) point. Statements containing many propositions received many points accordingly. The maximum scores possible for the African material were: 80.5 (prose), 81.0 (summary), and for the linguistic material: 77.75 (prose), 77.0 (summary). The point assignment was done independently by each of the authors and there was high agreement in assignment of points. Any discrepancies were resolved before the scoring began.

The statements in the subjects' recall protocols were compared with the statements in the corresponding summary or text versions. Only statements that had been underlined in the text version were scored in this initial analysis. Each protocol was scored by two scorers. Although the scoring was done independently, agreement of point assignment was quite high. Table 1 lists the mean number of points accumulated for each topic for each type of format; the separate ratings for each scorer are given along with the combined. The

Table 1

Mean number of idea points[a] scored for summary recall, underlined text recall, and total text recall as a function of topic and scorer.

	Original text				Summary	
	Underlined text		Total text			
	Scorer 1	Scorer 2	Scorer 1	Scorer 2	Scorer 1	Scorer 2
Linguistics	14.7	14.3	18.1	18.2	21.2	21.1
Mean	14.5		18.2		21.1	
African Geography	13.1	14.1	19.1	19.6	26.0	25.9
Mean	13.6		19.4		26.0	
	14.0		18.8		23.5	

[a] See text for scoring procedure.

high degree of consistency across the two scorers is an index of the reliability of the coding scheme.

As the summary statistics indicate, there was an extremely large effect of format, such that people remembered far more of the central ideas when they read the summary version than when they read the original text. The distribution of number of points across the two conditions were quite distinct. That is, the worst performance in the summary condition was better than most protocols in the text condition. Nonetheless, we performed a 2-way analysis of variance on the data. In addition to the principal factor of text versus summary format, we included factors of coder (to see if there was any difference or any interactions with the scorer), with topic (the linguistic versus African geography passage), and order of presentation of the material. The only significant effect was for type of format (i.e., text versus summary), $F(1, 12) = 22.9$, $p < 0.01$, such that summaries were much better than texts. No other effect or interaction had an F as large as 2.0.

Scoring the recall protocols from the original text version by only counting the underlined facts puts subjects at a disadvantage in that condition relative to recall in the summary condition. To the extent that subjects remember or try to recall other facts from the original text passage, these facts may be interfering with subjects' ability to recall the underlined facts. Therefore we did a second analysis where we counted any facts that were recalled from the original text, using the same point assignment system. It should be noted that in this situation subjects in the summary condition were at a disadvantage. In this case there were only one-fifth the potential facts to recall in the summary condition. The recall performance using this scoring technique is reflected in the Total Text column of table 1. As can be seen, performance is still considerably worse in the original text condition, 18.8 versus the 23.5 for summary.

A second analysis of variance was performed on these data using the same factors that were used above. Again, only the effect of information format reached significance, $F(1, 12) = 5.08$, $p < 0.05$, such that summaries were still superior. It is important to note that subjects were not instructed to recall facts that had not been underlined in the text, and therefore it is not surprising that they did not recall much more information than for just underlined facts. However, subjects were also told not to worry about whether a fact had been underlined or not (i.e., if they were unsure they should simply recall the fact). Therefore, since total recall was even worse than the amount of recall in the summary condition, it seems unlikely that poor performance in the text condition can be attributed to output interference.

Effects of format on clustering in recall

In addition to looking at amount recalled as a function of input format, we decided to measure the degree of organization of the recall protocols as

compared to the organization presented in the studied material. Both format types had used the same headings, subheadings, and numbering systems. If the original text included a specific subheading, the summary facts for that paragraph were preceded by the same section heading underlined. Since both formats had the same number of section headings, it was possible to use the Adjusted Ratio of Clustering (ARC) measure of clustering developed by Roenker et al. (1971). This clustering measure has the virtue of avoiding confounding the measure with absolute number of categories recalled, the distribution of items across categories, or the total number of items recalled.

The clustering measure, too, revealed an advantage for the summary condition, 55.65 versus 41.7 for text. This difference, however, did not reach statistical significance, $F(1, 12) = 3.94$, $p < 0.07$. Again, none of the other factors had an effect (e.g., scorer, topic) on clustering.

Discussion

The present study extends the results reported by Reder and Anderson (1980, 1982). Our results show that summaries give better retention than the original text even with a free recall test as the retention measure. Indeed, the advantage of summaries is even larger with this measure than with prior measures. The initial motivation for this line of research was to show the benefit of details on retention of important textbook ideas. We continue to be surprised at each study that supports the conclusion that details hurt retention of main points. These results are important pedagogically as they have implications for the optimal use of study time. Moreover, in one set of studies (Reder and Anderson 1982) we found that even when study time for main points is equated, there is no advantage for extra time spent on details. Not only was there no additional benefit of details, those details actually hurt retention of the main points even though they were not competing for study time with those main points.

The advantage of summaries over original texts can not be due totally to an inability for readers to isolate the main points from the background material that is of lesser importance. Although underlining the important points in a text reduced the advantage in prior studies, the advantage of summaries was still maintained in both true/false texts and in this experiment. Further, in a controlled laboratory experiment where main points and details were presented on a video screen, such that the main points were clearly distinguished from details, the advantage of the no detail (or summary) condition was still quite strong.

Two possibilities occur to us as explanations for this counterintuitive result. One possibility is that the details that are incorporated in textbooks are really poor embellishments of the material that is more central to the topic. This

possibility was argued extensively in Reder (in press a, b). It may be that elaborations generated by the comprehender are superior to those an author can provide in terms of mnemonic facilitation. There is evidence that elaborations are more effective when generated by the comprehender (Bobrow and Bower 1969; McFarland et al. 1980). Readers may be more inclined to generate elaborations in the summary condition than in the original text condition since they have more spare time.

The post-hoc analysis of clustering in recall provides another possible explanation for the advantage of the summary condition over the text condition in this experiment. For whatever reason, subjects' recall protocols were better organized or more clustered (as defined by the original input) when the information had been studied in summary format. (Of course, it is hard to determine whether better memory was due to higher clustering or higher clustering was due to better memory.) Probably the better organization or representation for information read in summary form is not the total explanation for summary's memory advantage. Although a good organization can definitely facilitate retrieval in a recall situation, the advantage of summary was even found in a true/false test where such a structural or organizational difference would presumably not apply. It should be reiterated, however, that the advantage of summaries was greater in this context. This means that the effects of organization or clustering may be important in facilitating the retention of summaries.

Conceivably the major advantage of the summary format is the more transparent organization that it affords. Once this organization is in memory, the more embellished text version might be assimilated with greater ease. There is evidence (Reder and Anderson 1980) that new information is learned better if prior related information was learned in summary form. Perhaps studying the same topic in text form after first studying it in summary form is the optimal way to spend one's time. That is, conceivably embellishments do facilitate retention of the main points if the appropriate organization or structure is first constructed in memory. An investigation of these ideas is currently underway.

References

Anderson, J.R. 1974. Retrieval of propositional information from long-term memory. Cognitive Psychology 5: 451–474.
Bobrow, S. and G.H. Bower. 1969. Comprehension and recall of sentences. Journal of Experimental Psychology 80: 455–461.
Bugelski, B.R. 1962. Presentation time, total time, and mediation in paired associate learning. Journal of Experimental Psychology 63: 409–412.
Cooper, E.H. and A.J. Pantle. 1967. The total-time hypothesis in verbal learning. Psychological Bulletin 68: 221–234.

Gleason, H.A. 1967. An introduction to descriptive linguistics. New York: Holt, Rinehart & Winston.

Hance, W.A. 1975. The geography of modern Africa. New York: Columbia University Press.

McFarland, C.E., Jr., T.J. Frey and D.D. Rhodes. 1980. Retrieval of internally versus externally generated words in episodic memory. Journal of Verbal Learning and Verbal Behavior 19: 210–225.

Postman, L. 1961. 'The present status of interference theory'. In: C.N. Cofer (ed.), Verbal learning and verbal behavior. New York: McGraw-Hill. pp. 152–179.

Reder, L.M. In press a. 'Techniques available to author, teacher and reader to improve retention of main ideas of a chapter'. In: J. Segal, S. Chipman and R. Glazer (eds.), Thinking and learning skills: current research and open questions, Vol. 2. Hillsdale, NJ: Erlbaum.

Reder, L.M. In press b. Elaborations: when do they help and when do they hurt? Text 2.

Reder, L.M. and J.R. Anderson. 1980. A comparison of texts and their summaries: memorial consequences. Journal of Verbal Learning and Verbal Behavior 19: 121–134.

Reder, L.M. and J.R. Anderson. 1982. Effects of spacing and embellishment on memory for the main points of a text. Memory and Cognition (in press).

Roenker, D.L., C.P. Thompson and S.C. Brown. 1971. Comparison of measures for the estimation of clustering in free recall. Psychological Bulletin 76: 45–48.

Smith, A.D. 1971. Output interference and organized recall from long-term memory. Journal of Verbal Learning and Verbal Behavior 10: 400–408.

Carl Martin Allwood and *Torbjorn Wikstrom* are presently at the University of Goteborg in the psychology department. They are engaged in study and research concerned with complex problem solving. They have also studied with H.A. Simon at Carnegie-Mellon.

Lynne M. Reder is an assistant professor of psychology at Carnegie-Mellon University, Pittsburgh, Pennsylvania. She received a B.A. from Stanford University in 1972, a Ph.D. from the University of Michigan in 1976, and spent two years at Yale as a post-doctoral fellow. She has written articles concerned with many aspects of memory, as well as papers on language processing.