

Very Long-Term Memory for Prose and Verse

DAVID C. RUBIN

Lawrence University

Recalls from five passages learned by undergraduates in the course of growing up in America were obtained. Unlike passages learned in the laboratory, the recalls, while partial, were exact with no evidence of constructive memory. Although there was no control over learning, practice or retention interval, the data are among the most regular in cognitive psychology. Function word, first letter, and music prompts increased recalls while they decreased a marked primacy effect evident in the free recall data. Free recalls obtained from fifth and sixth graders resembled the adult data. Recalls tended to begin and end at breath pause locations. The results fit a simple model of associative chaining retrieval of passively stored surface structure units.

While memory for discourse has been studied extensively under laboratory conditions, there are few systematic data on retention intervals longer than those convenient for laboratory study (Bahrick, Bahrick, & Wittlinger, 1975; Cofer, 1943b; Squire, Chace, & Slater, 1975; Titchener, 1923; Warrington & Sanders, 1971). In an attempt to provide such data, very long-term memory was tested for several prose and poetry passages that subjects were likely to learn in the course of growing up in America. Experiment 1 analyzes the free recall data of college students from three such passages. Experiment 2 uses the recalls of fifth and sixth graders who were more recently exposed to the material in order to obtain a shorter retention interval. Experiments 3 and 4 use various prompts to examine coding in more detail and to try to separate failures in retrieval from failures in retention.

This research was supported in part by NIMH Grant No. 1 RO3 MH27024-01 and in part by Lawrence University. I wish to thank Francis Campos, Cindy Marriott, Edwin Olson, and the principal and teachers of the Lincoln School, Appleton, Wisconsin, for their help. Address correspondence to David C. Rubin at Department of Psychology, Lawrence University, Appleton, Wisconsin 54911.

Copyright © 1977 by Academic Press, Inc.
All rights of reproduction in any form reserved.
Printed in Great Britain

EXPERIMENT 1

Method

Subjects. Ninety-two undergraduates volunteered.

Procedure. Covered booklets were used. On the top of each of the first three pages was one of the following titles: "The Preamble to the Constitution," "The 23rd Psalm: A Psalm of David," or "Hamlet's Soliloquy." The order of the titles was random. The last three pages were questionnaires which asked for each passage: "Did you have to memorize it?", "If so, when?", and "What experience have you had with it since?"

The subjects were asked to write down as much of the first titled passage as they could, and then to proceed through the booklet at their own pace.

Results

"The Preamble," "The 23rd Psalm," and "Hamlet's Soliloquy" were reported memorized by 32, 32, and 11 of the 85, 88, and 82 subjects who could answer the appropriate question. The average duration since memorizing the three passages was given by these subjects as 8 ($SD = 4.2$), 10 ($SD = 4.4$),

and 4 ($SD = 1.7$) years. Twenty-five, 19 and 18 subjects would report the time of their last experience with the passages. These times averaged 4 ($SD = 3.8$), 4 ($SD = 6.3$), and 3 ($SD = 1.9$) years ago. Thus, from the subjects' reports it appears that the recalls are from very long-term memory, and that less than half of the subjects memorized the passages. Discussions with local educators support these reports.

The recalls, while not complete, were verbatim: With few exceptions, subjects either recalled portions of the passage correctly in their original wording or not at all. While Bartlett's (1932) subjects reconstructed "The War of the Ghosts" after the passage of many years, the subjects of this study did not reconstruct the Constitution, the Bible, or Shakespeare.

In order to quantify this finding the recalls were scored on a word-by-word basis. Words that appeared more than once (for the most part, function words) had to appear in their proper context. Misspellings were the only change from verbatim recall allowed, and these only if they did not change the number or tense of the word. As many versions of "The 23rd Psalm" exist, the scoring of this passage had to be relaxed: If a word occurred in any of the several versions sampled, it was scored as correct on the King James version.

Of the 92 subjects tested 52, 53, and 60 were able to remember something of "The Preamble," "The 23rd Psalm," and Hamlet's Soliloquy. These subjects recalled an average of 14 ($SD = 12$), 42 ($SD = 32$), and 17 ($SD = 12$) correct words out of a possible 52, 117, and 277 words. They wrote down an average of 4 ($SD = 6$), 5 ($SD = 7$), and 3 ($SD = 7$) incorrect words. All remaining errors were errors of omission. Even using strict verbatim scoring the number of words in error divided by the total number of words recalled, or the error rate, is only about 15%. Thus, the scoring does not force the results, as only 15% of the words written do not enter into the analysis as correctly recalled units. To appreciate how different this finding is from laboratory recall of the results, a verbatim

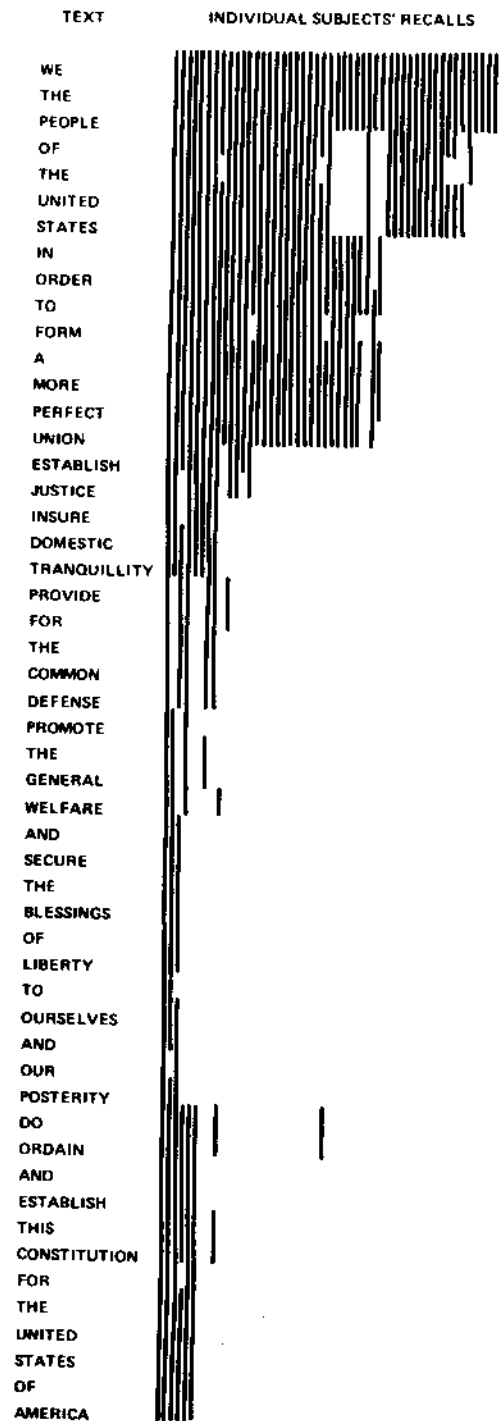


FIG. 1. Free recall data of 50 subjects for "The Preamble to the Constitution."

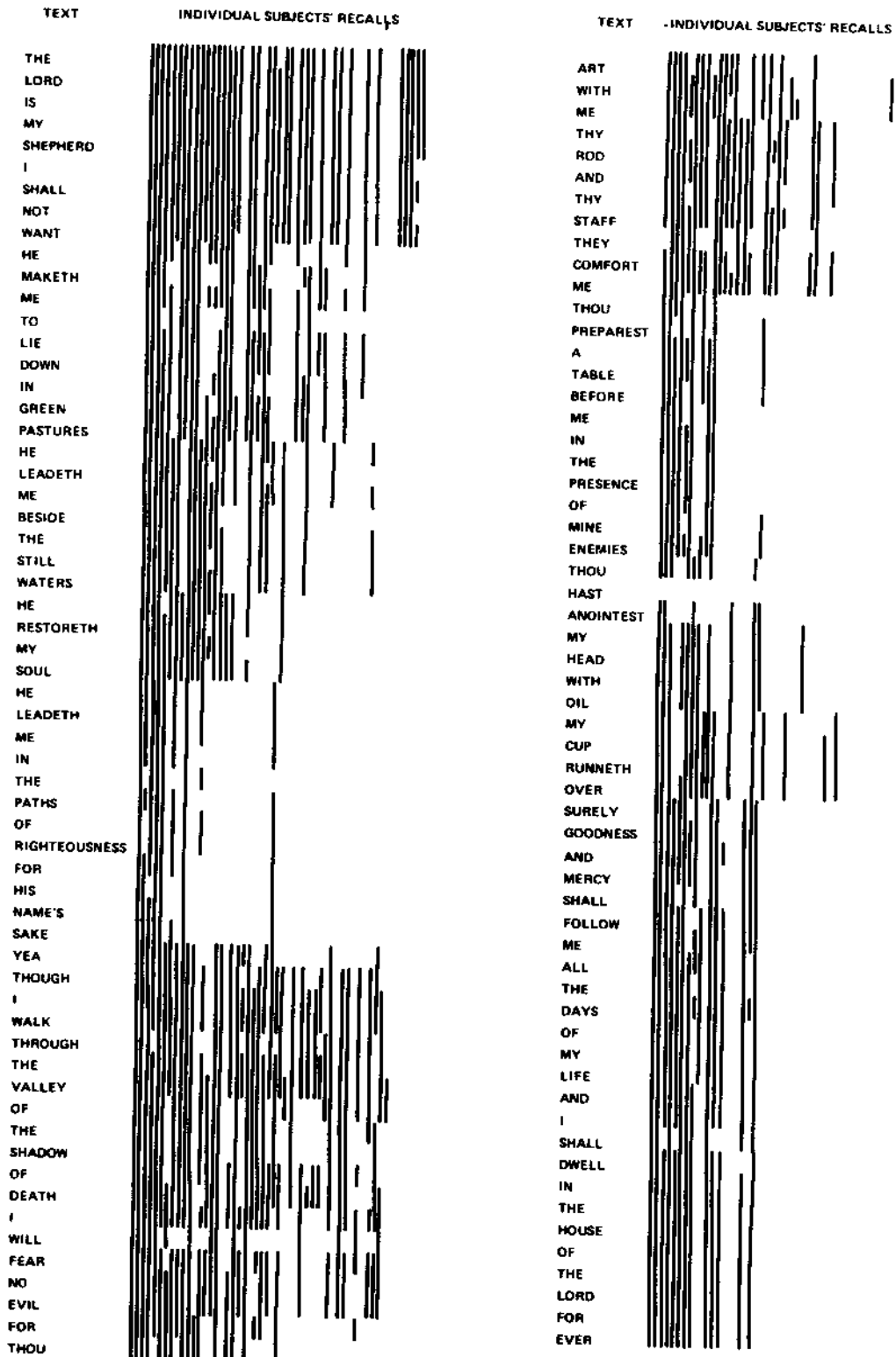


FIG. 2. Free recall data for "The 23rd Psalm."

scoring of Bartlett's "The War of the Ghosts" story was attempted. For the five initial free recalls for which Bartlett (1932, pp. 66-67) provides data, the average error rate is 69%. That is, using the strict verbatim scoring used here, Bartlett's subjects recalled over twice as many incorrect words as correct words, whereas the subjects in the present study recalled less than one-fifth as many incorrect words as correct words.

Besides being accurate, the recalls are regular across subjects. Figures 1, 2, and 3 each display the recall data for 50 individual subjects selected randomly from among those who recalled something from the passage listed. Each of the 50 vertical columns represents one subject: A line indicates that the word was correctly recalled by that subject,

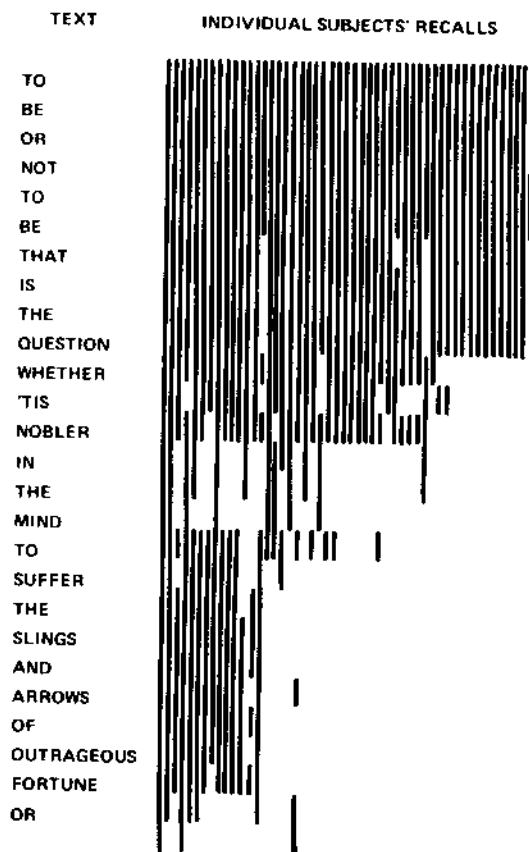


FIG. 3. Free recall data for those words of "Hamlet's Soliloquy" recalled by four or more subjects.

and a blank indicates that it was not. Hamlet's rather long soliloquy was stopped where fewer than four of the subjects remembered any of the remaining words. It should be stressed that each of these figures contains data for 50 individual subjects. The only manipulation of the raw scores that has been made is that the subjects are rank ordered by the amount they recalled. The recalls of the subjects who remembered the most are placed farthest to the left.

The data are quite regular. The Guttman Coefficients of Reproducibility for the "The Preamble," "The 23rd Psalm," and "Hamlet's Soliloquy" are .950, .832, and .991. This implies that if the number of words an individual subject remembers is known, exactly what words that subject remembers can, on the average, be predicted with an accuracy of 95, 83, and 99% from the number of times each word was remembered by the group as a whole. That is, if a subject recalled 20 words correctly, the prediction that they would be the 20 words most often recalled by the group as a whole would be correct 95, 83, and 99% of the time for the three passages. Thus, each individual is representative of the group. While it is apparent from the figures that the data are far from random, an exact statistical by-chance measure is difficult with as many subjects and words as are used here (Kenny and Rubin, 1977). While qualitatively similar findings are present in laboratory studies (Rubin, 1974), considerably less regularity is found.

For "The Preamble" and "Hamlet's Soliloquy," primacy provides a good description of the data. Based on primacy alone an individual subject's data for the three passages, can be predicted with 92.9, 76.1, and 98.8% accuracy. Here, as primacy is a prediction that is independent of the data, a binomial by-chance model (Kenny and Rubin, 1977) is appropriate and yields significant differences: 69.4%, $t(49) = 11.17$; 69.3%, $t(49) = 4.18$; and 88.8%, $t(49) = 12.53$; all $p < .001$. Again, it should be stressed that these are predictions of which words individual subjects will recall. These are not predictions of grouped data but,

rather, the average of predictions of individual subjects' data. The fact that the primacy effect is not as strong for "The 23rd Psalm," which has a simple rhythmic structure, will be investigated in more detail after examining other passages.

Qualitative examination of Figures 1 to 3 shows that subjects tend to begin and end recalling at surface structure boundaries corresponding to what might be breath pause locations (Glanzer, 1976; Suci, 1967). In order to quantify this notion, operational definitions of boundaries in recall and of breath pause locations are needed.

A recall boundary was defined as (1) more than two consecutive words correctly recalled followed by more than two consecutive words not correctly recalled: that is, an end of a portion of recall, or (2) more than two consecutive words not correctly recalled followed by more than two consecutive words correctly recalled: that is, a beginning of a portion of recall. One consequence of this definition is that a recall boundary cannot occur in the first or last three words of a passage, so the observation that almost all recalls begin with the first word of the passage is not used to inflate the correspondence of breath pauses and recall boundaries.

Johnson's (1970) technique was followed to quantify breath pauses. Five volunteers enrolled in a Psychology of Language course were asked to "please indicate by means of slashes any location where a pause would be acceptable (i.e., word/word)." Judgements were made on all three passages used here as well as two others used in later experiments. A breath pause was defined as any location that three or more of the five judges marked.

"The Preamble," "The 23rd Psalm," and "Hamlet's Soliloquy" consist of 52, 117, and 277 words, respectively. They therefore have 51, 116, and 276 boundaries between words. Of these, 39, 38, and 43% are breath pause locations. Of the 64, 151, and 62 recall boundaries observed for the three passages, 94, 81, and 82% fall on the breath pause locations. As some of the subjects provide more than one re-

call boundary an inferential statistical analysis of this data is difficult. However, if only the first recall boundary provided by each subject is considered, the observations can be assumed to be independent and a χ^2 analysis is appropriate. Of the 45, 46, or 40 first recall boundaries, 93, 87, and 80% fall on breath pause locations: $\chi^2(1) = 53.08$, $\chi^2(1) = 44.97$, and $\chi^2(1) = 21.12$, all $p < .001$. The 5, 4, and 10 subjects whose recalls yielded no recall boundaries were by necessity omitted from this analysis. Thus, the recall boundaries coincide with independently determined surface structure units.

Discussion

Given no control over learning conditions, practice, or retention interval, the recall data produced are among the most regular in cognitive psychology. Not only do all subjects tend to remember the same words, differing only in the amount they recall, but counter to all current thinking about memory for discourse, they also remember in a non-abstractive, nonreconstructive manner. The recall is accurate and organized in terms of surface structure units. In terms of depth of processing approach, there is evidence of only the shallowest processing. Thus, in an attempt to extend laboratory findings to the real world, interesting limitations have been encountered.

EXPERIMENT 2

Anecdotal evidence and psychological theory suggest that all very long-term memory is not verbatim. The three passages sampled as being likely to be learned by college sophomores in the course of growing up are also likely to be among those that have their exact wording held in high regard. Thus, it is likely that the demand characteristics of recalling this type of material, rather than the long retention intervals, are responsible for the results. To exclude the very long retention interval as a factor, children who had recently been exposed to similar material for the first time were tested.

Method

Subjects. Fifty-one fifth graders and 43 sixth graders who had learned about, but had not been required to memorize, "The Preamble to the Constitution" and Lincoln's "Gettysburg Address" were tested.

Procedure. Approximately 6 months after learning about the two passages, half of the fifth graders were verbally asked to write down "The Preamble"; when they had finished they were asked to write down the "Gettysburg Address". The other half were given the titles in the reverse order. This procedure was repeated for the sixth graders approximately 18 months after they had learned about the passages. There was no formal practice of the two passages in the school during the retention intervals.

Results

Of the 51 fifth graders, 18 were able to recall something of "The Preamble," and 16 were able to recall something of the "Gettysburg Address." These subjects recalled an average of 14.2 ($SD = 4.2$) and 4.9 ($SD = 2.1$) correct words and 1.3 ($SD = 1.0$) and 0.0 ($SD = 0.0$) incorrect words from the two passages. Of the 43 sixth graders, 13 were able to recall something of "The Preamble" and 17 were able to recall something of the "Gettysburg Address." These subjects recalled an average of 13.2 ($SD = 10.6$) and 8.7 ($SD = 4.7$) correct words and .8 ($SD = .7$) and .4 ($SD = .8$) incorrect words from the two passages. As the fifth and sixth grades had different teachers at the time they were originally exposed to the passages, and according to these teachers different learning experiences, a direct comparison between the amount recalled by the two groups would be difficult to interpret.

The Guttman Coefficients of Reproducibility are: .962 for the fifth graders' "Preamble" and .999 for the fifth graders' "Gettysburg Address," .938 for the sixth graders' "Preamble" and .991 for the sixth graders' "Gettysburg Address." The corresponding values using primacy instead of

empirically determined rank order are only slightly lower: .942, .998, .911, and .990. Under a binomial model the by-chance values would be .616, .963, .700, and .937, respectively. Again primacy is a statistically significant predictor: $t(17) = 13.93$, $t(15) = 9.28$, $t(12) = 5.14$, and $t(16) = 7.48$, all $p < .001$.

In "The Preamble" 39% of the boundaries between words are breath pause locations; in the "Gettysburg Address" 37% are. Of the fifth graders' 17 and 13 recall boundaries on these two passages, 88 and 69% are at pause locations. Of the sixth graders' 15 and 14 recall boundaries, 93 and 43% are at pause locations. Considering only the first recall boundary provided for each subject, in order to meet the independence assumption of a χ^2 test, 92% of the 12 fifth graders' "Preamble" recall boundaries, 69% of the 13 fifth graders' "Gettysburg Address" recall boundaries, 90% of the 10 sixth graders' "Preamble" recall boundaries, and 43% of the 14 sixth graders' "Gettysburg Address" recall boundaries were on breath pause locations: $\chi^2(1) = 11.74$, $p < .001$; $\chi^2(1) = 4.57$, $p < .05$; $\chi^2(1) = 8.80$, $p < .01$; $\chi^2(1) = 0.04$, n.s.

Thus, the fifth and sixth graders' data are quite similar to the adults' in accuracy of recall, scalability, primary, and unit boundaries. A more direct test is possible. Using the rank ordering of words from most to least likely to be recalled from the adult "Preamble" data of Experiment 1 and from the adult "Gettysburg Address" data from Experiment 3, the recall of individual children can be predicted with an accuracy of 95.5, 99.2, 89.9, and 98.6%. As the rank ordering of words is determined independently of the data being tested, the binomial model used to test the primacy effect provides a valid test of scalability: $t(17) = 14.85$, $t(15) = 9.22$, $t(12) = 5.52$, and $t(16) = 9.34$; all $p < .001$.

One case of reconstructive memory was noted in the Gettysburg Address. Of the 14 subjects who recalled the word "fathers," six added the prefix or word "fore." This error, which in violation of the initial scoring rules was counted as being correct, fits in with both

the archaic wording and repeated mention of the syllable "fore" in the first sentence: "Fourscore and seven years ago our fathers brought forth ..." Even this reconstruction, however, appears to be capable of being explained on the surface structure level.

Discussion

As a longitudinal design was not used, the fifth and sixth graders differ from the undergraduates on many dimensions. This coupled with the lack of controlled learning conditions makes estimation of the rate of decay impossible. In spite of these differences, however, the elementary school students and the undergraduates tend to show marked similarity in their recall: both in the lack of reconstruction and in the actual words they recall. It thus appears that there is something about the way the passages are learned or recalled, rather than the very long retention interval, that produces the anomalous recall.

EXPERIMENT 3

If the recall is really in surface structure units, then surface structure cues should improve recall. Such prompts should also be able to assess to what extent the failure in recall is a failure of retrieval rather than of retention.

Method

Subjects. Twenty-six undergraduates volunteered for the experiment.

Procedure. The procedure was the same for all subjects. On the first page of a booklet subjects were asked to write down as much as they could of the "Gettysburg Address." When all subjects had finished they were asked to turn to a version of the "Gettysburg Address" which had all content words (Fries, 1952) replaced by blanks of uniform length. Thus, the passage began "____ and ____ _____ our _____ on this ____ a ____" Again they were asked to write down all that they could putting only one word in each blank. Finally, the subjects were

presented with the form they had just seen with first letters inserted (e.g., "f____ and s____y____a____..."). Subjects then filled in a questionnaire.

While no control group totally unfamiliar with the "Gettysburg Address" was available, it seems highly unlikely that many of the missing content words could have been guessed correctly from the cues given, unless the subjects were familiar with the "Gettysburg Address."

Results

Of the 26 subjects, 9 had memorized the "Gettysburg Address" at an average of 11 ($SD = 4.6$) years ago.

The prompts increased recall, $F(2, 50) = 71.23, p < .001, MS_e = 77.17$, with an average of 7.5, 10.9, and 26.4% of the content words correct in the free recall, context, and context plus first letter conditions, respectively. Thus, there appears to be more information in memory than can be retrieved in free recall.

The data, by necessity scored for only content words, were again quite orderly with Guttman Coefficients of Reproducibility of .973, .951, and .893 for the three conditions. The predictive value of primacy was only slightly less than the Coefficient of Reproducibility but was reduced by the prompts, predicting 97.2, 94.1, and 78.8% of the responses correctly. Thus, the primacy phenomena can be attributed in part to retrieval effects. The by-chance binomial model predicts 86.7, 81.4, and 66.7% of the responses correctly, again significantly different from primacy: $t(25) = 7.01, t(25) = 7.97, t(25) = 9.37$, all $p < .001$.

To allow comparisons with the previous two experiments the free recall data was also scored for all words as well for content words only. For the 22 subjects who recalled something in the free recall condition, the average number of correct words was 20 ($SD = 12$) and the average number of incorrect words was 4 ($SD = 3$). As with the fifth graders, a large proportion of the subjects who recalled the word "fathers" recalled it as "forefathers"

(12 out of 22). Again this substitution was considered as correct. The Guttman Coefficient of Reproducibility is .976. Primacy would predict 97.0% of the correct responses, and the by-chance binomial model would predict 86.7%, which is statistically different from primacy: $t(21) = 8.35, p < .001$. Of the 266 boundaries between words in the "Gettysburg Address," 37% are breath pause locations. Of the 31 free recall boundaries, 90% fall at breath pause locations. For the 20 first recall boundaries, 95% are at breath pause locations: $\chi^2(1) = 26.68, p < .001$. Thus, the free recall data are in general agreement with that of Experiments 1 and 2.

EXPERIMENT 4

Except for "The 23rd Psalm," which has a simple rhythmic structure, the data from previous experiments show a very strong primacy effect. In order to include a second passage of known rhythmic structure and in order to demonstrate that recalls from such passages are remembered, in part, by the use of rhythmic structures, subjects were asked to recall "The Star Spangled Banner" either with no music, "The Star Spangled Banner", or "Stars and Stripes Forever" playing. If the subjects' recall units were basically rhythmic, then the proper rhythm should increase the amount recalled and the wrong rhythm should decrease the amount recalled (Glanzer, 1976).

Method

Subjects. Ninety-five undergraduates volunteered for the experiment.

Procedure. Tape recordings were made of a continuously repeating 1 min, 10 sec instrumental version of "The Star Spangled Banner" and a 2 min, 45 sec instrumental version of "Stars and Stripes Forever." "The Star Spangled Banner" was played to one group of subjects, "Stars and Stripes Forever" was played to a second group, and nothing was played to a third group. All three groups were asked to "please write down the

National Anthem, that is, the words to the music that is played at the beginning of sporting events." A questionnaire followed.

Results

There were 31, 32, and 32 subjects in the right, wrong, and no music conditions. Twelve, 17, and 10 of these subjects reported having memorized "The Star Spangled Banner" out of the 28, 30, and 28 subjects answering the question. The average duration since last exposure was 12 ($SD = 4.0$), 11 ($SD = 2.5$), and 8 ($SD = 3.2$) years for the 12, 17, and 7 subjects able to answer that question. There were no significant differences between the groups on these questions.

The strongest results cannot be given here in a quantitative fashion. The subjects in the no music condition behaved as most subjects in verbal learning experiments do, with perhaps more of an expression of frustration. The subjects in the wrong music condition appeared to be in slightly more pain, and a few on occasion put their hands over their ears. Most interesting, however, were the subjects in the right music condition. They would write as fast as they could until the music got ahead of them, and then they would switch to the behaviors of the wrong music condition. By the second or third repetition of "The Star Spangled Banner" almost all of the subjects in the right music condition adopted a strategy of waiting until the music came around to where they had stopped writing the previous time, and then writing another burst until the music got ahead of them again. The effect was quite striking and has since provided an effective classroom demonstration of the role of coding in memory.

The subjects in the right, wrong, and no music conditions remembered an average of 52, 28, and 32 of the 80 words of "The Star Spangled Banner" correctly: $F(2, 92) = 12.68, p < .001, MS_e = 422$. Their recall was quite accurate with an average of only 3.4 ($SD = 2.7$), 1.6 ($SD = 1.8$), and 2.5 ($SD = 2.2$) incorrect words recalled.

As in previous experiments, the data were quite orderly with Guttman Coefficients of Reproducibility of .814, .921, and .892. Under the binomial chance model these figures would be .632, .702, and .658. For the wrong and no music conditions primacy was again a fairly good predictor of which items would be recalled with coefficients of .836, $t(31) = 4.91$, $p < .001$, and .852, $t(31) = 6.31$, $p < .001$. However, for the right music condition primacy was essentially equivalent to chance: .624, $t(30) = 0.39$, n.s. This is not due to a ceiling effect as, even in this condition, subjects recalled only 65% of the words correctly.

As "The Star Spangled Banner" consists of 80 words there are 79 between word boundaries, of these 38% were acceptable breath pauses. Of the 80, 47, and 35 recall boundaries found in the right, wrong, and no music conditions, 58, 77, and 71% fell on breath pauses. Of the 27, 27, and 23 first recall boundaries, 59, 78, and 70% fell on breath pauses: $\chi^2(1) = 4.33$, $\chi^2(1) = 16.75$, $\chi^2(1) = 8.45$, all $p < .05$.

Discussion

Recalls from college students for five passages have now been presented. Three of these passages ("The Preamble," "Hamlet's Soliloquy," and the "Gettysburg Address") have no simple, repetitive, rhythmic structure to aid in recall. For these three passages primacy predicts free recall with an accuracy of 93, 99, and 97%. Two of the passages, "The 23rd Psalm" and "The Star Spangled Banner," have marked rhythmic structures. For these two passages primacy predicts free recall with an accuracy of 76 and 85%. These results and the raw recall data, such as are shown in Figures 1 to 3, are consistent with the following hypothesis. All five passages are remembered by associative chaining of surface structure units. Thus, if one unit is forgotten all remaining units are lost. Where a rhythmic structure is available, subjects can make use of this structure to begin remembering the chain again after a unit is forgotten. While this

hypothesis is obviously inadequate for all serial learning (Lashley, 1951; Jones, 1974), it provides a good account of the data collected here. It may be noted that this hypothesis does not make use of the most commonly cited determinant of prose memory: meaning. This is not because meaning has definitely been shown to be irrelevant to the present study, but rather because meaning would greatly increase the complexity of the hypothesis without increasing its ability to account for the data.

The prompted recall studies provide information about whether this chaining is a function of retrieval or of storage. When prompts are available, such as the function words for the "Gettysburg Address" in Experiment 3 or in the correct music for "The Star Spangled Banner" in this experiment, the primacy effect is reduced. Thus, it appears that the chaining is at least partially a function of the retrieval process. In order to provide a quantitative assessment of the extent of the loss in primacy as prompts are provided, correction must be made for the changes that also occur in the chance level prediction. Green (1956) provides such a corrected index called the Index of Consistency. The amount by which the coefficient is greater than its chance level is divided by the total amount that it could be greater than its chance level, or in the present case: $(\text{Primacy} - \text{chance}) / (1.00 - \text{chance})$. The Index of Consistency has expected values between 1 (complete scalability) and 0 (scalability equal to chance). For the "Gettysburg Address" with no prompts, content word prompts, and content word plus first letter prompts, the Indexes of Consistency are .789, .683, and .363, respectively. For "The Star Spangled Banner" with the wrong music, no music, and the correct music, they are .450, .567, and -.022, respectively. As a large portion of the primacy effect in the "Gettysburg Address" experiment and all of the primacy effect in "The Star Spangled Banner" experiment can be attributed to retrieval phenomena, it appears that the chaining is mostly due to retrieval strategies.

GENERAL DISCUSSION

A sample of prose and poetry consisting of "The Preamble," "The 23rd Psalm," "Hamlet's Soliloquy," the "Gettysburg Address," and "The Star Spangled Banner" produced extremely regular free recalls although no control over learning conditions, practice, or retention interval was available. The evidence for verbatim, nonconstructive recall, surface structure units of breath pause length, and extremely simple chaining and rhythmic structures found in the free recall data was supported by prompted recall studies. The results, however, while internally consistent, are counter to almost every study of memory for discourse and almost all theoretical views. How is this difference to be reconciled? Two possibilities will be considered.

The first possible reconciliation is to claim that, as the five passages sampled are "good figures" in the Gestalt sense, they show little change in recall. That is, as the passages are already in their most stable state, there is little reason for them to change in transmission through one person. The five passages are surely among the best written in the English language, yet their exact wording is far from what you would expect from an undergraduate or a fifth grader. That is, it is unlikely that even the best writings of a hundred or more years ago would still be "good figures" today.

A second possible reconciliation is argued by Neisser (1967) under the term of "demand characteristics." People normally remember stories in a given way partly because of the limits on their cognitive capacity and partly because of what is expected of them. When someone asks you to "read a story through twice at your normal reading speed and write down what you remember as accurately as you can," not only could you not write down the whole story, but also you would interpret the question to mean write down the gist as accurately as you can. A different expectation comes if the passage is from the Constitution, the Bible, or Shakespeare. Neisser's argument

is, in part, that just because subjects behave one way under a given set of expectations they do not always have to behave that way. Memory for discourse in our culture is generally reconstructive, but it does not have to be.

How the demand characteristics actually function through initial learning and later recall has not been studied here in any detail. To do this properly would have required controlled learning conditions and subjects who were initially naive both to the passages and to their sacred nature. Some speculation, however, is possible. As one-third of the subjects reported memorizing the passages while two-thirds reported not memorizing them, it is unlikely that any kind of purposeful memorization—incidental learning distinction could account for the differences between the results obtained here and other studies of memory for discourse. A more likely alternative is to assume that the passages were learned and recalled with emphasis on their exact wording and possibly even without understanding of their overall meaning. If this is the case, then the literature on remembering discourse verbatim and on remembering discourse where there is little internal structure to aid in learning and recall would be relevant.

The analyses performed on most past studies make comparisons of error rates, scalability, and surface structure units difficult. Several studies using either verbatim learning or passages with minimal structure have, however, analyzed serial position effects. If the process operating in these studies are similar to the ones operating here, then a marked primacy effect would be expected. Unfortunately, this is not always the case. Using sentences and a word-by-word serial anticipation method, Mandler and Mandler found a serial position effect based on "the core meaning of the sentence" (1964, p. 197) rather than on primacy. Using passages with so little internal structure that the order of sentences could be randomized, both a primacy and a smaller, but reliable, recency effect were found

(Deese & Kaufman, 1957; Jersild, 1929; Russell & Sewall, 1972).

A more direct study also argues against explaining the anomalous results obtained here in terms of the distinction between learning passages for their exact wording and learning them for their gist. Cofer (1941, 1943a) examined, in great detail, the effect of instructions to learn the exact wording of passages versus instructions to learn their essential ideas. While more trials were required to learn passages to criterion under the exact wording condition, Cofer (1943a) concluded that for the shape of the learning curves and the types of errors obtained, there was little observable difference due to the instructions.

Thus, the question of the way in which the demand characteristics actually function through learning and recall does not seem to be answered in any obvious manner by the dimensions of memorization—incidental learning and verbatim—substance learning studied in the laboratory. Whatever the relevant processes of learning and recall are, however, the regularity in the data indicates that they are the same for all subjects.

REFERENCES

- BAHRICK, H. P., BAHRICK, P. O., & WITTLINGER, R. P. Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, 1975, **104**, 54–75.
- BARTLETT, F. C. *Remembering: A study in experimental and social psychology*. London: Cambridge University Press, 1967. (Originally published, 1932).
- COFER, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. *American Journal of Psychology*, 1941, **54**, 1–20.
- COFER, C. N. An analysis of errors made in the learning of prose materials. *Journal of Experimental Psychology*, 1943, **32**, 399–410. (a)
- COFER, C. N. Recall of verbal material after a four-year interval. *Journal of General Psychology*, 1943, **29**, 155–156. (b)
- DEESE, J., & KAUFMAN, R. A. Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, 1957, **54**, 180–187.
- FRIES, C. C. *The structure of English*. New York: Harcourt, Brace, and World, 1952.
- GLANZER, M. Intonation grouping and related words in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 85–92.
- GREEN, B. F. A method of scalogram analysis using summary statistics. *Psychometrika*, 1956, **21**, 79–88.
- JERSILD, A. Primacy, recency, frequency and vividness. *Journal of Experimental Psychology*, 1929, **12**, 58–70.
- JOHNSON, R. E. Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 12–20.
- JONES, M. R. Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.
- KENNY, D. A., & RUBIN, D. C. Estimating chance reproducibility in Guttman scaling. *Social Science Research*, 1977, **6**, 188–196.
- LASHLEY, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium*. New York: Wiley, 1951.
- MANDLER, G., & MANDLER, J. M. Serial position effects in sentences. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 195–202.
- NEISSER, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- RUBIN, D. C. *Remembering prose*. Unpublished doctoral dissertation, Harvard University, 1974.
- RUSSELL, G. W., & SEWALL, E. C. Serial-position effect and organization of recall of connected meaningful verbal material. *Psychological Reports*, 1972, **30**, 443–446.
- SQUIRE, L. R., CHACE, P. M., & SLATER, P. C. Assessment of memory for remote events. *Psychological Reports*, 1975, **37**, 223–234.
- SUCL, G. J. The validity of pause as an index of units in language. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 26–32.
- TITCHENER, E. B. Relearning after forty-six years. *American Journal of Psychology*, 1923, **34**, 468–469.
- WARRINGTON, E. K., & SANDERS, H. I. The fate of old memories. *Quarterly Journal of Experimental Psychology*, 1971, **23**, 432–442.

(Received January 24, 1977)