Episodic and Semantic Aspects of Memory for Prose

D. James Dooling and Robert E. Christiaansen
Kent State University

Subjects' knowledge about the topic of a prose passage was manipulated to evaluate the locus of constructive processes in remembering. Subjects who knew during comprehension that a passage was about a famous person falsely recognized thematic sentences at 2-day and 1-week retention intervals. Subjects told during the retention interval of the famous main character's identity made thematic errors only after 1 week. The main character manipulation precipitated the most thematic errors when it came immediately after comprehension. The results are interpreted in terms of the episodic-semantic memory distinction. With the passage of time, subjects have increased difficulty in retrieving passage-specific episodes. They compensate by using related information from semantic memory.

According to Bartlett (1932), remembering prose is a constructive process. Meaningful material is stored in memory in schematic form and recall is achieved by a process of reconstruction. Bartlett observed that recall becomes increasingly abstract, thematic, and distorted with the passage of time. That is, at long intervals, there is less recall of passage-specific information and more recall from the subjects' general knowledge of the topic. The present study deals with constructive remembering over time. We attempt to understand prose retention in terms that are compatible with Bartlett's conception of remembering.

One problem for a constructive theory of remembering is that subjects do recall some of the exact words that they have experienced in an experiment (Cofer, 1973). A complete theory of prose retention, therefore, must contain mechanisms that account for both reproductive and constructive aspects of remembering. In the present study, we use Tulving's (1972) distinction between episodic and semantic memory as a framework for conceptualizing changes in remembering with the passage of time.

Tulving observed that words from an experimental list are encoded contextually. They are represented in memory as episodes that occurred at a particular time and place in the life of the rememberer. Episodic memory is sharply distinguished from semantic memory, which is conceptually organized. It is the storehouse of the subjects' preexperimental knowledge. Tulving (1972) did not comment on prose retention, but remembering a prose passage surely involves both the episodic and semantic memory systems. Preexperimental knowledge is used in comprehension, but subjects also encode various aspects of the text episodically. With the passage of time, the retrieval of episodic information becomes increasingly difficult, so subjects supplement their memory performance with related knowledge. Constructive memory for prose, therefore, is seen as a filling in of episodic gaps with knowledge from semantic memory.

The experiments reported here follow up on a study previously published by Sulin and Dooling (1974). In that experiment, the subjects' knowledge of a topic was manipulated by giving them a passage.
that could be about either a famous or a fictitious person. For example, one group of subjects read a short biography of a ruthless dictator named Gerald Martin. A second group read the same passage, but with the name of the main character changed to Adolph Hitler. Retention was measured on a recognition test for sentences from the passage intermixed with foil sentences. At a short, 5-min. retention interval there were few false recognitions of foils thematically related to the famous person. But at a 1-week interval such errors increased dramatically for subjects who had read about the famous main character. For example, those who had read the Hitler version tended to falsely recognize a sentence referring to his persecution of the Jews. The result replicated Bartlett's (1932) observation that thematic intrusion errors increase with the passage of time.

The present study attempts to extend the Sulin and Dooling (1974) results by focusing on the locus of the constructive effects that were obtained. Cofer (1977) points out that constructive processes can occur at three places in the sequence between input and test: (a) during initial comprehension; (b) during the retention interval; (c) during the retention test. Most research on constructive memory to date has focused on the effect of comprehension processes (e.g., Bransford & Johnson, 1972; Dooling & Lachman, 1971; Frederiksen, 1975). But changes in memory performance with the passage of time, such as were found by Sulin and Dooling (1974), seem to imply that important constructive activity occurs subsequent to initial encoding. In the present study, therefore, we attempt to manipulate postcomprehension constructive activity and to assess its importance for an understanding of prose retention. Experiment 1 focuses on a manipulation that occurs just before the retention test after a short versus long interval. Experiment 2 studies the effect of such a cognitive manipulation on constructive processes at two different points within the retention interval. Both studies deal with the use of subjects' knowledge from semantic memory to compensate for their inability to retrieve episodic information from the passage itself.

Experiment 1

Our strategy was to test retention of the Sulin and Dooling (1974) materials over two intervals, manipulating when the subjects learned that the passage was about a famous rather than a fictitious person. In addition to the two conditions used in the previous study (famous vs. fictitious), the key experimental group for the present research was the "after" condition. These subjects read a fictitious passage at input but found out the true identity of the famous person just before the recognition test. Thematic errors by the after group would reflect the use of semantic memory at output, not constructive encoding of the material at input. As episodic information is forgotten with the passage of time, we expected the thematic performance of the after group to approach that for the famous condition.

There was one other experimental condition, the "before" group. They were included to control for the fact that the after group had to switch between a fictitious and famous main character between input and test. The before group read the same materials as the after group (the fictitious version of a passage) and were also informed that the material was actually about a particular famous person. Unlike the after group, however, the before group was given the main character instructions before they read the passage. We expected the before group to perform the same as the famous group. If the after group performed differently from the famous group, the difference could not be attributed to generalized confusion from the manipulation of the main character's identity.

Method

Subjects. The subjects were 320 introductory psychology students who were fulfilling a course requirement. Fourteen additional subjects were run at the 2-day retention interval because 10 subjects
did not show up for the retention test and four experimenter errors were committed in the distribution of test materials. At the 1-week interval, 7 additional subjects were run. Five failed to report for the test, and 2 were involved in procedural mistakes.

**Materials.** The two passages from Sulin and Dooling (1974) were used. One is about a ruthless dictator who could either be the fictitious person, Gerald Martin, or Adolph Hitler. The other is about a troubled young girl whose fictitious name was Carol Harris. In the famous condition, Helen Keller was the main character. Each passage was 10 sentences long and written in a style such that subjects could not guess the famous main character of the fictitious passages. The passages can be found in the previous publication (Sulin & Dooling, 1974, p. 256).

The recognition test materials were also the same as those used by Sulin and Dooling. Seven "old" sentences were selected from the appropriate passage. There were also seven new or "foil" sentences that varied in thematic relatedness. Four were taken from the passage not read and were thematically neutral or nonthematic. The other three varied in thematic relatedness at low, medium, or high levels. The three thematic new sentences presented true facts about the famous person, but they were not sentences from the passage itself. For example, the high-thematic foil sentence for the Harris/Keller passage was: *She was deaf, dumb, and blind*. Sulin and Dooling (1974) provide rating information on the thematic relatedness of the foil sentences (p. 261).

All sentences for both input and test were printed on the blank side of an IBM card. The passages for input were presented as a deck of 10 cards. The recognition test consisted of a randomly mixed deck of 14 cards.

**Procedure.** The subjects were run in groups of 10 or less with the four major conditions evenly represented within a session wherever possible. Subjects were assigned to conditions by haphazardly distributing the input decks to various desks and allowing subjects to choose their own seats.

Tape-recorded instructions informed all subjects that they would be reading a passage once and that their memory would be tested later. They were instructed not to discuss the passage with anyone during the retention interval and told to return for a test at the appropriate time, either 2 days or 1 week later.

On the test day, all subjects first read a printed set of instructions. The famous, before, and fictitious groups read the same instructions that they had received before input. But the after group was given a new instruction that was analogous to that received by the before group on Day 1: "The passage that you read in the last session was an abbreviated biography of a fictitious person named Gerald Martin (Carol Harris). But the passage was actually about a famous person named Adolph Hitler (Helen Keller). That is, the name 'Gerald Martin' (Carol Harris) was substituted for Adolph Hitler's (Helen Keller's) name throughout the passage." All subjects were given 1 min. to think about the passage prior to the recognition test.

The recognition test consisted of seven old and seven new sentences randomly intermixed. Subjects were informed that about half of the sentences were old. They sorted the cards in a self-paced manner into one of six slots: The three slots on the left were labeled "yes" for the old sentences and the three on the right were labeled "no." Within each category, there were three degrees of confidence: low, medium, and high.

**Design.** The design was a 4 x 2 x 2 factorial with 20 subjects in each cell. The variables were (a) main character—famous, before, after, fictitious; (b) interval—2 days, 1 week; (c) passage—Hitler/Martin, Keller/Harris. For the major analysis of performance on new sentences there was an additional variable that was within subjects: (d) foil type—neutral, low-, medium-, high-thematic.

**Results.** Scoring. Separate analyses were performed on the data from the old and new sentences. For the old sentence analysis, each subject was given a percent correct score based on the number of yes responses for the seven sentences that had occurred in the passage. For the new sentence analysis, each subject was given four scores to represent four levels of thematic relatedness: neutral, low, medium, and high. Yes responses were scored as 1 and no responses as 0. Performance on the four neutral foils was averaged to yield a single neutral score and each subject was given either a 1 or 0 for each of the low-, medium-, and high-thematic foils. The means from the analysis on this measure are expressed as a per-
percentage of false-positive errors on each foil type. We also performed an analysis of variance on arcsin transformations of the proportion data and on combination scores that weight error with confidence to yield a 6-point scale. Since all three analyses yielded essentially the same conclusions, we have reported only the simple percents. In all analyses, the .05 level of significance was used.

New sentences. The means of central interest are those reflecting the false-positive errors for the various foil types, as shown in Figure 1. The solid lines replicate the famous versus fictitious conditions from the Sulin and Dooling (1974) experiment. The greater divergence of the solid lines at the 1-week interval replicates the previous finding that thematic effects increase with the passage of time. For the present experiment, a three-between, one-within mixed analysis of variance was conducted. All main effects were significant: main character, \( F(3, 304) = 7.30, MS_e = .16 \); interval, \( F(1, 304) = 4.86, MS_e = .16 \); passage, \( F(1, 304) = 7.62, MS_e = .16 \); and foil type, \( F(3, 912) = 35.76, MS_e = .09 \).

The means for three of these effects can be inferred from the graph. In addition, there was a greater proportion of false-positive errors with the Hitler/Martin passage than with the Keller/Harris passage (.18 vs. .12).

Figure 1 shows that the after group performed like the fictitious group at the shorter interval, but more like the famous and before groups at the longer interval. The triple interaction shown, however, was not statistically significant. But a more focused analysis on critical points in the graph is more relevant for the questions asked. Since our main concern was with differential performance with thematic material, we conducted planned comparisons for the high-thematic foil data only. The pooled error term was .11, and the degrees of freedom have been adjusted following Winer (1971, p. 530). At 2 days, the after and fictitious conditions were not significantly different; but at 1 week, the after group made significantly more false recognitions of the high-thematic foil, \( F(1, 554) = 7.41 \). Comparison of after performance with the before group showed them making significantly fewer high-thematic errors at the 2-day interval, \( F(1, 554) = 5.67 \). At 1 week, there was no significant difference for the before-after comparison. For high-thematic material, the after group showed a selective increase in false-positive errors with the passage of time.

The analysis of variance yielded two significant interactions. Foil type interacted with passage, \( F(3, 912) = 24.84, MS_e = .09 \). The interaction was mainly due to unusual performance with the medium-thematic foil for the Hitler/Martin passage. As happened in the Sulin and
Dooling (1974) experiment, there were many false recognitions of this particular foil across all main character conditions. When averaged with the Keller/Harris data, the medium-thematic foil yielded the pronounced elevation that is evident in the graph. Our major conclusions are not, however, dependent on performance with the medium-thematic foils. The other significant interaction was that of Main Character X Foil Type, $F(9, 912) = 4.02$, $MS_e = .09$. This interaction was expected, as we had anticipated that the differences among the main character conditions would increase with increasing thematicity of the foils.

Old sentences. Analysis of variance was performed on the correct yes responses (hits) for the seven sentences that had actually occurred in the passage. Subjects were fairly accurate at correctly recognizing old sentences, 82% at 2 days and 80% at 1 week. None of the effects in the analysis of variance was significant. The fairly high accuracy does not necessarily imply that the subjects had an excellent episodic memory for individual sentences from the passage. The old sentences are also thematic to the passage and subjects may have responded on that basis. In general, the analysis of hits tells us little about the nature of the memory processes under investigation.

Discussion

Performance of the after group in Experiment 1 demonstrated constructive memory performance that was not entirely determined by the conditions of initial encoding. The after group read the same materials with the same instructions as the fictitious group, yet they made significantly more high-thematic errors at the 1-week interval. Some might argue that we have demonstrated a response bias on the part of our subjects. Our interpretation is different, however. From our perspective, the after group engaged in a constructive retrieval process that is an integral part of remembering. It is particularly noteworthy that the thematic effects changed with the retention interval. The after manipulation did not simply introduce a generalized response tendency on the part of the subjects. They responded thematically only after they had forgotten much of the episodic information from the passage. When subjects were unable to retrieve the specific memory codes that they had previously stored, they used related knowledge from semantic memory to fill in the gaps. Our results, therefore, are consistent with Bartlett’s (1932) theory of constructive remembering.

Experiment 2

In Experiment 1, the important manipulation came right before the retention test. The constructive processing that occurred outside of initial comprehension, therefore, occurred at the time of the retention test. The present experiment investigated the interaction between the after manipulation and processes during the retention interval. Experiment 1 showed that the constructive effects in the after condition were dependent on forgetting during a 1-week retention interval. Experiment 2 held the retention interval constant at 1 week but varied when the after manipulation occurred. We added a group of subjects who were told the name of the famous main character immediately after comprehension. The famous condition was dropped, as the before condition was sufficient for demonstrating thematic effects. All four groups of subjects, therefore, were presented with the fictitious version of a passage. Three of the groups were told that the passage was actually about the famous person, either before processing it, immediately after processing it, or a week after processing it. Performance of these three groups was compared to that for the fictitious group that dealt with a fictitious passage for both study and test.

Because three of the experimental conditions replicated Experiment 1, the predictions for them were straightforward. The before versus fictitious groups should differ substantially in false recognitions of thematic foils. The week-after group, which was the same as the after group from
Experiment 1, should make significantly more high-thematic errors than the fictitious group. Predictions were less obvious for the immediately-after group. If the after manipulation were interpreted as introducing a response bias, then the week-after and immediately-after groups should not differ. But we predicted that the after manipulation would interact with processes occurring during the retention interval. Subjects who had integrated a fictitious passage with a famous main character 1 week before the retention test should have a more difficult time discriminating between the passage that they read and their general knowledge of the topic. It was predicted, therefore, that the immediately-after group would make more thematic errors than the week-after group.

Method

Subjects. The subjects were 216 students of introductory psychology who were fulfilling a course requirement. Seventeen additional subjects were run: Six failed to show up for the memory test, and 11 were involved in some form of procedural error.

Materials. The passages and foils were exactly the same as in Experiment 1. For input, each fictitious passage was tape recorded in a normal tone of voice. For the recognition test, the 14 old and new sentences were intermixed on a single sheet of paper in one of two random orders. The randomization was subject to the constraints that the high-thematic sentence appear in the middle of the page (Position 7 or 8) and that there never be more than 3 old or new sentences in a row.

Procedure. Subjects were run in large groups in a classroom. We tried to have about 15 subjects in each group, but the group sizes actually ranged from 14 to 19. The instructions were substantially the same as in Experiment 1. For input, all fictitious passages were tape recorded in a normal tone of voice. For the recognition test, the 14 old and new sentences were intermixed on a single sheet of paper in one of two random orders. The randomization was subject to the constraints that the high-thematic sentence appear in the middle of the page (Position 7 or 8) and that there never be more than 3 old or new sentences in a row.

Prior to input, all groups were given an instruction about the passage to be heard: Three groups were told that they would be listening to a fictitious biography, and the before group was instructed about the famous person. Then the tape of the passage was played. At this point, the immediately-after group was given the famous-main-character instruction followed by a 1-min. rehearsal period. All groups were sent home for 1 week with instructions not to discuss the material with anyone. When they reported for the retention test, all groups were given an instruction about the passage they had read the previous week.

Three of the groups were told that it had actually been about the famous main character. This information was new only to the week-after group. The fictitious group was given a repeat of their previous instruction. One minute of silent rehearsal was then given before the self-paced recognition test. Subjects checked off yes or no for each of 14 sentences and rated their confidence for each decision on a 3-point scale.

Results

New sentences. The percentage of false-positive errors for the foil sentences is shown in Figure 2, collapsed across the passage variable. Analysis of variance was performed on both percentage of errors and yes-confidence combination scores. Since there was little difference between the two, we report only the percentage of false-positive errors. The three main effects in the analysis were significant: main character, $F(3, 208) = 8.09$, $MS_e = .21$; passage, $F(1, 208) = 24.77$, $MS_e = .21$;
foil type, $F(3, 624) = 29.96, MSe = .11$. In addition, the Passage $\times$ Foil Type interaction was significant, $F(3, 624) = 21.87, MSe = .11$. Once again, the medium-thematic foil for the Hitler/Martin passage yielded an exceptionally high proportion of false-positive errors.

The one other significant effect reflected the result of major importance, the Main Character $\times$ Foil Type interaction, $F(9, 624) = 2.35, MSe = .11$. The means are shown in Figure 2. Three aspects of the results should be considered: (a) The two solid lines, before versus fictitious, substantially replicated the Sulin and Dooling (1974) results and comparable results from Experiment 1 of the present study. (b) The week-after data were very similar to those obtained in the present Experiment 1. On the high-thematic foil, for example, the week-after group made significantly more false-positive errors than the fictitious group, $F(1, 762) = 7.13, MSe = .13$, but did not differ significantly from the before group. (c) The immediately-after group made the most false-positive errors on thematic foils. Considering only the high-thematic foil, as we have done in other comparisons, the difference did not reach statistical significance. But an analysis that combined the three thematic foils did show a significant difference between the immediately-after and the week-after groups, $F(1, 762) = 16.19, MSe = .13$. We had predicted that the immediately-after group would make more errors than the week-after group. But we did not anticipate that they would make significantly more thematic errors than the before group, $F(1, 762) = 9.56, MSe = .13$.

Old sentences. Analysis of variance on the seven old sentences from the passage yielded no significant effects. The mean percent correct was 84.

Discussion

Experiment 2 had two after conditions, both of which involved the same main-character manipulation between comprehension and test. The immediately-after and week-after groups did, however, differ significantly in thematic performance. The effects demonstrate an interaction between constructive memory processing and cognitive events occurring within the retention interval. Subjects' knowledge of the topic was manipulated in a way that affected their constructive memory retrieval processes. Such processes change with the quality of the memory codes available at different points during the retention interval. The results point out the importance of constructive memory processes that occur after initial comprehension of the material. To consider such processes to be indicative of a "response bias" is to dismiss by fiat some very interesting aspects of remembering.

We did not predict that the immediately-after group would make more thematic errors than the before group. But they did. In additional research performed in our laboratory, Shutts (Note 1) also obtained a similar effect. We are confident, therefore, that the results can be replicated for the type of manipulations and materials we have used. If replicated in other laboratories, the results could prove to be important, as some of the best studies of constructive memory have obtained their effects with an immediately-after manipulation (Loftus, 1975; Spiro, 1977). More recently, however, Loftus, Miller, and Burns (Note 2) have reported results somewhat different from ours. They found that misleading information had a greater effect on memory distortion when it occurred just before the retention test, rather than immediately after comprehension. There are many differences between that experiment and the present one, so the source of the discrepancy is difficult to pin down. Our task seems to require an active cognitive reorganization of the full content of the material when the name of the famous main character is given. The Loftus manipulation, on the other hand, focuses on one particular detail of the material to be remembered. With two such different memory tasks for the subjects, differential effects are quite understandable. Clearly, we need to know a great deal more about
the effects of manipulating knowledge during the retention interval.

General Discussion

The results of the two experiments taken together support a constructive view of remembering as suggested by Bartlett (1932). Subjects at a 1-week retention interval have lost much of the specific information from the passage. In Tulving’s (1972) terms, they have an inability to retrieve such information from episodic memory. Under such conditions, it is natural for them to use their general knowledge, their semantic memory, to guide their test performance. They do not do so at random. Rather, they make their retention performance fit with general ideas that they remember from the passage. Thus, Bartlett’s well-known finding that recall becomes more thematic with the passage of time was replicated here with a recognition technique. As the interval lengthens, memory for passage-specific episodes is lost, and test performance reflects increasing amounts of output from semantic memory.

Of particular interest in the present study was the manipulation of the subjects’ knowledge during the retention interval. In our after conditions, subjects read a passage about a fictitious person and were told later that it had actually been about a particular famous person. At that point, they had new knowledge, a new schema, that could affect their memory for the material. It is our view that they attempted to integrate that new knowledge with the information acquired at input. The result was thematic memory performance. The effect cannot be attributed to a simple response bias or the demand characteristics of the experiment. Thematic performance for the after groups depended on their cognitive state during the retention interval. In Experiment 1, they did not make thematic errors at the shorter interval. They still had good memory for the experimental material and could use that episodic knowledge to edit their recognition responses. Constructive performance was obtained only after a 1-week interval. Royer, Perkins, and Konold (Note 3) also failed to find thematic effects for an after group at a short retention interval. In Experiment 2, maximum thematic performance was obtained for subjects who had to perform a difficult cognitive integration immediately after reading the passage. In a sense, our immediately-after subjects were in a position very similar to those reading Bartlett’s (1932) War of the Ghosts or the Dooling and Lachman (1971) Christopher Columbus story. Subjects had to perform a difficult cognitive manipulation to comprehend the material. Their use of extra knowledge from semantic memory showed up a week later in thematic recognitions. Subjects in the week-after condition, however, did not have as much to integrate, as they had lost much of the specific memory for the passage during the retention interval.

A constructive theory of memory emphasizes the role of extraexperimental knowledge at both initial encoding and subsequent retrieval. The relative importance of constructive encoding and constructive retrieval depends to an important extent on the length of the retention interval and the events that occur during that interval. For example, Dooling and Mullet (1973) failed to show thematic effects for a manipulation that came immediately after comprehension. We now believe that constructive retrieval effects would have been obtained if the interval had been longer. Many constructive memory experiments in the literature are similar to the Dooling and Mullet study in using a very short retention interval. As more realistic retention intervals are investigated, we expect to see much more interest in constructive retrieval processes.

We have found Tulving’s (1972) distinction between episodic and semantic memory useful for describing the remembering of prose over long intervals. A constructive theory of memory focuses on the use of a subject’s semantic memory in the storage and retrieval of episodes. Clearly, some distinction between preexperimental knowledge and experimental material is needed.
We do not, however, wish to claim agreement with every detail of Tulving's (1972) formulation. Tulving's conception of episodic memory was developed on the basis of word-list experiments in which subjects were scored for the correct recall or recognition of exact words. For us, however, the episodes comprehended from prose exist at various levels of abstraction. Subjects remember not only some of the exact words, but also ideas at various levels. Constructive memory retrieval depends on an episodic memory for some abstract representation of the text, a concept that has not been considered by Tulving. A constructive theory of remembering needs to take into account the fact that high-level concepts are remembered longer than low-level concepts (Dooling & Christiaansen, 1977). The more abstract the concepts available at the time of the retention test, the greater the role of semantic memory will be in constructive recall.

Reference Notes

2. Loftus, E. F., Miller, D. G., & Burns, H. J. Integration of verbal information at various times following a visual event. Paper presented at the meeting of the Psychonomic Society, St. Louis, Missouri, November 1976.

References


Received January 31, 1977