INTRUSION OF A THEMATIC IDEA IN RETENTION OF PROSE

REBECCA A. SULIN AND D. JAMES DOOLING

Kent State University

Two implications of Bartlett's theory of memory were tested: (a) that prose passages are stored in schematic form, and (b) that thematic assimilation increases with the passage of time. The subjects read brief biographical passages about either a famous or a fictitious person (e.g., Adolph Hitler vs. Gerald Martin). Recognition memory for individual sentences was tested after intervals of either 5 min. or 1 wk. As expected, passages with a famous main character yielded more false positive errors. In addition, the errors in this condition depended on the thematic relatedness of the recognition foil. The further prediction on thematic assimilation was also upheld: Thematic effects were relatively greater at the longer retention interval.

The resurgence of cognitive psychology, as defined by Neisser's (1967) classic text, has led to a revival of interest in the work of Bartlett (1932) on remembering. According to Bartlett, the central meaning of a prose passage is stored in memory in schematic form. Recall of the words is achieved by a process of reconstruction from this underlying theme. Since memory is hypothesized to be an active reconstructive process in which Ss make use of their knowledge of the world, thematic intrusion errors should be a common feature of recall. For example, in Bartlett's "War of the Ghosts" story, two Indians are "hunting seals." During recall, S might remember that the Indians were "fishing." The general idea was remembered, but S's knowledge of the world intrudes with the more culturally common interpretation of the event to be remembered. With longer retention intervals, Bartlett reports that such thematic tendencies become more pronounced. As the schema becomes more dominant, specific details are assimilated.

More recently, Pompi and Lachman (1967) have demonstrated that Ss make thematic errors in the retention of prose. They had Ss read a highly thematic passage about an operation for cancer. The individual words of the story, however, were low associates to the theme. For example, the surgeon was "Chief Resident Jones" who uncovered an "ugly growth too large for removal." On a subsequent recognition test, Ss tended to falsely recognize highly thematic words that had not actually occurred in the passage (e.g., surgeon, cancer, etc.). A similar effect has been obtained by Bransford and Franks (1971). After their Ss had read a series of related sentences, they tended to falsely recognize new sentences that contained ideas integrated from those previously presented. Both of these experiments show that memory is "constructive" in Bartlett's sense of the term.

The present study attempts to extend our knowledge of constructive memory processes by providing a manipulation of the schematic knowledge used by Ss in remembering. The Ss were asked to read a short biographical passage. For half of them, the main character was a famous person, like Adolph Hitler. In reading the passage, these Ss could draw on their pre-experimental knowledge of Hitler to aid them in the comprehension and retention of the passage. The other half of the Ss were asked to read the same passage, but presented as a biography of a fictitious main
character named Gerald Martin. It was assumed that Ss who read the famous main character version of a passage would be working with a more richly elaborated schema. They would, in effect, call into their working memory a good deal of extraneous information that was not specifically required by the passage to be remembered. These Ss, therefore, would be more likely to make a false positive error on a subsequent recognition test.

A second variable in the experiment had to do with the characteristics of a particular recognition foil, hereafter referred to as the key sentence. The key sentence was "new," it was presented in the recognition test but had not occurred in the original passage. For half of the Ss, the key sentence was high in thematic relatedness. It asserted a well-known fact that was part of the schematic stereotype of the famous main character. The other half of the Ss were tested with a low-thematic key sentence, one that described a less dominant attribute of the famous person. This manipulation of the key sentence, along the dimension of thematic relatedness, allows us to assess whether false recognitions are specifically thematic, as would be predicted by Bartlett's theory of the schema. The Ss who read about the famous main character should make more false positive errors on the high- than on the low-thematic key sentence. No such prediction is made for the fictitious main character condition.

A third independent variable was the length of the retention interval. The recognition test was administered after either 5 min. or 1 wk. Every theory of memory predicts poorer retention at the longer interval. Bartlett's theory, however, specifically requires that thematic errors become relatively more pronounced with the passage of time. Specific information is expected to decay with time, while the schema remains relatively stable and intact. With the passage of time, Ss should rely more on the general information contained in the schema, and less on the specific information remembered from the passage. If Ss rely more heavily on the schema at the longer retention interval, they should make more highly thematic false recognitions. In short, the thematic interaction predicted with the main character and key sentence variables should be of greater magnitude after 1 wk. than after 5 min.

Two experiments are reported. They are nearly identical in conceptual design, but differ in a number of methodological details.

EXPERIMENT I

Method

Design. A $2 \times 2 \times 2 \times 2$ factorial design was employed with 10 Ss in each cell. The factors were: (a) main character—famous vs. fictitious; (b) key sentence tested—high vs. low thematic; (c) interval—5 min. vs. 1 wk.; and (d) passage—Adolph Hitler (Gerald Martin) vs. Helen Keller (Carol Harris).

Subjects. The Ss were 160 introductory psychology students who were fulfilling a course requirement. Three additional Ss were run in the 1-wk. condition, but had to be replaced because they did not show up for the retention test.

Materials. Two short biographical passages were constructed by the authors to meet the requirements of the experimental manipulations. Each passage was written so that its assertions were true of a specific famous historical character. Individual sentences, however, were fairly low in thematic relatedness when taken out of context. The passages were pretested informally and rewritten until naive readers (a) agreed that the sentences were true of the intended famous main character, and (b) failed to recognize the famous main character of the story when the name of a fictitious character was substituted. The text of Passage 1 is given below:

Helen Keller's need for professional help. Helen Keller was a problem child from birth. She was wild, stubborn, and violent. By the time Helen turned eight, she was still unmanageable. Her parents were very concerned about her mental health. There was no good institution for her problem in her state. Her parents finally decided to take some action. They hired a private teacher for Helen.

A second version of this passage was created by substituting the name Carol Harris for Helen Keller throughout.

The other experimental passage was about Adolph Hitler. It was written in the same style and also consisted of a title and seven sentences. The fictitious character for the Hitler story was named Gerald Martin. Passage 2 is given below:

Adolph Hitler's seizure of power. Adolph Hitler strove to undermine the existing government to satisfy his political ambitions. Many of the people of his country supported his efforts. Current political problems made it relatively easy for Hitler to take over. Certain groups
remained loyal to the old government and caused Hitler trouble. He confronted these groups directly and so silenced them. He became a ruthless, uncontrollable dictator. The ultimate effect of his rule was the downfall of his country.

For each passage, a highly thematic key sentence was also constructed. The key sentence was an alternate version of one of the sentences in the original passage. It had the same grammatical structure as the original and fit appropriately into the context of the passage. Whereas the original sentence was fairly low in thematic relatedness, the high-thematic key sentence asserted a well-known central fact about the famous main character. The content of this sentence was determined by the ratings of 34 college students. They were given the name of a famous character and asked to write down eight sentences about his or her life. They were then asked to rank order each. For Helen Keller, the central fact was clearly that She was deaf, dumb, and blind. This sentence, therefore, was chosen as the high-thematic alternate for the low-thematic key sentence italicized in the passage above. The Adolph Hitler ratings were not as clear-cut. The dominant rating centered around the fact that he was the Nazi leader of Germany in World War II. This was not used, for it could not be plausibly attributed to Gerald Martin. Instead, the second most common response, having to do with his persecution of the Jews, was used: He hated the Jews particularly and so persecuted them. Further ratings on the thematic relatedness of sentences are given in the Results section of Experiment II.

Procedure. The Ss were run in groups of eight or less, with each S serving in a different experimental condition. Assignment to conditions was haphazard in that materials were distributed to eight desks and Ss were allowed to choose their own seats. The Ss were told that the experiment was about memory and that they would be reading a short story, one sentence at a time. Each sentence appeared on a separate page of a booklet. The Ss were instructed to read each sentence for 4 sec. and to turn the page in time with a click from E. The instructions emphasized reading for meaning and discouraged rote memorization.

After reading the story, half of the Ss were dismissed and asked to return at the same time 1 wk. later. The remaining Ss sat quietly in their seats for 5 min. Both groups were under instructions not to rehearse the material or to discuss it with anyone. After the appropriate retention interval, the recognition test sheets were distributed. Each contained seven of the eight original sentences in the correct order. The key sentence was switched for the recognition test so that Ss received the alternate of the one they had actually read. The Ss were instructed to read each sentence in order and to judge whether it was exactly the same, slightly different, or very different from the sentence originally read. They also rated their confidence for each judgment on a 5-point scale.

Results and Discussion

Scoring. The dependent variable for all analyses of Experiment I was a score that combined correctness of judgment with degree of confidence. Each of the three response categories was considered to have a range of 5 points, depending on S's confidence. For the key sentence, the correct judgment was very different. Within this category, the higher S's confidence, the higher his recognition score. At the opposite extreme, a response of exactly the same was in error. The higher S's confidence within this category, the lower his recognition score. Judgments of slightly different yielded intermediate scores. The scoring procedure is outlined in Table 1.

Recognition scores had a range of 15, from -7 to +7. All scores in the slightly different category were assigned zero values because the directionality of confidence was indeterminate. The scoring of non-key sentences followed the same procedure, but the directionality was reversed. For these, exactly the same was the correct response and yielded the higher scores. In both analyses to be reported below, therefore, high scores reflect high recognition accuracy.
Key sentence. The key sentence was the one and only new sentence on the recognition test for all Ss. Analysis of variance for recognition accuracy of this sentence yielded the predicted main effect of main character, $F(1, 144) = 13.21$, $p < .001$. The Ss who read a passage about a famous main character had lower scores than did those who read about a fictitious character (4.00 vs. 5.85). The main effect of interval was also significant, $F(1, 144) = 10.51$, $p < .01$. Performance was better after a retention interval of 5 min. than after 1 wk. (5.75 vs. 4.10). Neither of the two remaining main effects, type of key sentence and passage, was significant.

The predicted Main Character $\times$ Type of Key Sentence interaction was significant, $F(1, 144) = 5.10$, $p < .05$. The means for this interaction can be inferred from those presented in Table 2. As predicted, in the famous main character condition, performance was worse when the key sentence tested was high in thematic relatedness. This result is consistent with Bartlett's (1932) theory that Ss store a passage in schematic form, relying on their preexperimental knowledge of the topic as part of the remembering process. Because of schematic storage, a high-thematic key sentence presents S with a more difficult discrimination on the recognition test. Hence, there are more false positive errors. In contrast, Ss in the fictitious group, tested on the same material, give the opposite pattern of results. Recognition accuracy is actually higher on the high-

### Table 2

<table>
<thead>
<tr>
<th>Key sentence tested</th>
<th>Main character</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Famous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five-minute interval</td>
<td>Low thematic</td>
<td>5.45 (.80)</td>
<td>5.95 (.90)</td>
</tr>
<tr>
<td></td>
<td>High thematic</td>
<td>4.70 (.80)</td>
<td>6.90 (1.00)</td>
</tr>
<tr>
<td>One-week interval</td>
<td>Low thematic</td>
<td>3.85 (.75)</td>
<td>4.75 (.80)</td>
</tr>
<tr>
<td></td>
<td>High thematic</td>
<td>2.00 (.50)</td>
<td>5.80 (.95)</td>
</tr>
</tbody>
</table>

Note. Proportion correct is included in parentheses.
central idea. The hypothesized organizational properties of a schema, therefore, were not effectively manipulated in the present study.

There was one significant interaction, Interval X Passage, $F (1, 144) = 4.15, p < .05$. The passage that was remembered better at 5 min. underwent a greater amount of forgetting during the 1-wk. interval. Neither of the significant main effects, however, was compromised by the interaction.

**Experiment II**

Two of the three predictions made in Experiment I were confirmed and the third was in the predicted direction. Experiment II was a replication of the first study with a number of minor methodological changes. On the recognition test, 5s were presented with seven old sentences intermingled with seven new sentences in random order. The decision on each sentence was a *yes-no* judgment that had to be made on each sentence individually and out of context. One of the seven recognition foils was the alternate form of the key sentence, as in Experiment I. The design was the same, with 11 5s in each of the 16 cells.

**Method**

**Subjects.** The 5s were 176 students at Kent State University who were fulfilling a course requirement. Eighty-five of these 5s were from the same population as those in Experiment I. Fifty-nine were summer school students of introductory psychology. Twenty were enrolled in a graduate level psychology course for nonmajors. Two because their recognition test booklets were defective, and 2 because English was not their native language.

**Materials.** The passages to be read were exactly the same as in Experiment I. For the recognition test, the seven non-key old sentences were presented along with seven new, foil sentences. One of these was the alternate version of the key sentence, just as it was in Experiment I. The six additional foils contained two that were constructed to be thematically related to the famous main character. It was thought that the presence of these two thematic foils would attenuate any novelty associated with the key sentence. In addition, these two sentences were chosen to represent differing degrees of thematic relatedness. One of the additional foils was constructed to be of medium thematic relatedness, and one was relatively low in thematic content. The ratings collected in Experiment I were used for this purpose. The low-thematic foil incorpo-

**Results and Discussion**

**Scoring.** The recognition score was a combination of accuracy and confidence, as in Experiment I. Correct scores were given a value from +1 to +5, depending on degree of confidence. The higher the confidence, the higher the recognition score. Erroneous judgments yielded negative recognition scores, with high confidence yielding the lowest score of $-5$.

**Key sentence.** Recognition scores on the key sentence alone were submitted to analysis of variance, as in Experiment I. There were significant main effects for both main character, $F (1, 160) = 10.3, p < .01$; and interval, $F (1, 160) = 27.7, p < .001$. The 5s were less accurate when the main character had been famous rather than fictitious (2.80 vs. 4.05), and performance was better at the shorter retention interval (4.44 vs. 2.40). The other two main effects were not significant. All main effects replicate the findings of Experiment
I. Unlike Experiment I, however, there was a significant Main Character X Interval interaction, $F(1, 160) = 5.5, p < .05$. The means can be inferred from those presented in Table 3. As predicted, the main character effect is of greater magnitude at the longer retention interval.

The predicted Main Character X Type of Key Sentence interaction, which was obtained in Experiment I, did not reach statistical significance at the .05 level, but these two variables interacted significantly with interval, $F(1, 160) = 5.7, p < .05$. The means are shown in Table 3. The predicted thematic effect was obtained only at the 1-wk. retention interval. The Ss who read the main character version of a passage had a greater tendency to falsely recognize a high-thematic key sentence. This result confirms Bartlett's (1932) contention that thematic assimilation increases with the passage of time. At the longer retention interval, Ss have less memory for the specifics of the passage and rely on reconstruction from the schema to a greater degree.

**Foil sentences.** The inclusion of seven recognition foils in Experiment II makes it possible to test the experimental predictions with a somewhat different design. Four of the foils were neutral with respect to theme, and these were averaged into a single score for each S. In addition, the present analysis includes data only from those Ss who were tested on the high-thematic version of the key sentence. When these two things are done, foil type can be viewed as a within-Ss variable with four levels: neutral and low, medium, and high thematic. This variable was combined factorially with the between-Ss variables of main character, interval, and passage for analysis of variance. It was predicted that recognition accuracy on the foils would depend on thematic relatedness and that the thematic effects would be greater at the longer retention interval.

In the analysis of variance, all of the main effects were significant beyond the .01 level of significance. The means for the effects of main character, $F(1, 80) = 10.9$, interval, $F(1, 80) = 21.0$, and foil type, $F(3, 240) = 10.7$, can be inferred from the data presented in Figure 1. As for passage, $F(1, 80) = 11.0$, performance was better on the Helen Keller/Carol Harris passage (4.31 vs. 3.31). The three possible two-way interactions among the variables of main character, interval, and foil type were all significant, but each of them was compromised by their triple interaction, $F(3, 240) = 4.7, p < .01$. The means for this interaction are shown in Figure 1. At the 5-min. retention interval, all means are fairly close to the maximum value of 5. There are no differences for either the main character or foil type variables. After 1 wk., however, Ss in the famous main character condition clearly perform less accurately. With this one group the type of foil has a pronounced effect. The more thematic the foil, the more likely it is to be falsely recognized. This interaction occurred in similar form for both experimental passages. The four-way interaction did not approach statistical significance.

There was one other significant interaction in the present analysis, that of Passage X Foil Type, $F(3, 240) = 6.4, p < .001$. The interaction was primarily due to the fact that performance was unusually low for the medium-thematic foil for the Adolph Hitler/Gerald Martin passage. The sentence in question was Hitler (Martin) was obsessed with the desire to conquer the world. There appeared to be a thematic effect on this sentence with both the famous and fictitious main characters. It is, evidently, a sentence that can plau-

---

**TABLE 3**

<table>
<thead>
<tr>
<th>Key sentence tested</th>
<th>Main character</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Famous</td>
</tr>
<tr>
<td>Five-minute interval</td>
<td></td>
</tr>
<tr>
<td>Low thematic</td>
<td>3.68 (.86)</td>
</tr>
<tr>
<td>High thematic</td>
<td>4.86 (1.00)</td>
</tr>
<tr>
<td>One-week interval</td>
<td></td>
</tr>
<tr>
<td>Low thematic</td>
<td>1.86 (.77)</td>
</tr>
<tr>
<td>High thematic</td>
<td>.77 (.59)</td>
</tr>
</tbody>
</table>

*Note.* Proportion correct is included in parentheses.
Possibly be applied to any dictator of the sort described in the passage. Performance on this one sentence is reflected in the one data point that appears to be out of place in Figure 1.

As a double check on the conclusions to be reached on the basis of the foil type analysis, a post hoc rating of thematic relatedness was performed by 20 introductory psychology students. The raters were given the 15 sentences associated with the famous main character version of each passage (7 old sentences, 2 key sentences, 6 non-key foils) and asked to rate each on a 7-point scale. They were told the name of the famous main character and asked to rate how central each sentence was with respect to the person's life. A rating of 7 was to be given only to sentences that asserted "the first thing someone would say in describing the person's life." A rating of 1 was reserved for "sentences that do not apply to the famous person or that could be about almost anyone." Two different random orderings of sentences were used for each passage and ratings of the two passages were done in counterbalanced order for equal numbers of raters. The results of the ratings support the previous classification of the foil types. The mean ratings for the neutral and low-, medium-, and high-thematic foil types were: (a) Helen Keller—1.68, 4.95, 6.15, and 7.00; and (b) Adolph Hitler—1.53, 4.65, 6.30, and 6.45. The scale along the abscissa of Figure 1, therefore, has ordinal but not interval properties. The 7 old sentences were rated 4.62 and 5.58 for the Keller and Hitler passages respectively, while the low-thematic version of the key sentence was rated 5.15 and 3.75. The latter rating is not relevant for the foil-type analysis, but it does confirm the effectiveness of the key sentence manipulation in the two previous analyses.

Old sentences. Recognition scores on the seven sentences that had appeared in both study and test were averaged for each S and analysis of variance was performed on this measure. The main effect of main character was significant, $F (1, 160) = 5.6$, $p < .05$. Performance was worse with the famous main character (2.17 vs. 2.76). The main effect of interval was also significant, $F (1, 160) = 7.17$, $p < .01$. Performance was better at 5 min. than at 1 wk. (2.80 vs. 2.14). The other two main effects were nonsignificant. There was only one significant interaction, that of Main Character X Passage, $F (1, 160) = 3.8$, $p < .05$. The difference due to the main character variable was in the same
direction for both passages, but it was of greater magnitude with the Helen Keller/Carol Harris passage.

GENERAL DISCUSSION

The results of the two experiments provide support for Bartlett's (1932) theory of memory. It was shown that preexperimental knowledge of the topic of a passage affects performance on a subsequent retention test. Specifically, false recognitions of thematically related foils show that memory for prose involves a reconstructive process from an abstract representation of the substance of the passage. A number of recent studies have also demonstrated the constructive nature of memory and substantially confirm Bartlett's insights on the remembering process (cf. Cofer, 1973). Of special relevance in the present study is the demonstration that thematic effects increase with the passage of time. At the longer retention interval, Ss have less memory for specifics and rely to a greater extent on their memory for the theme. Retention interval effects have been almost totally ignored by students of constructive memory processes. The present results suggest, however, that forgetting from long-term memory can be profitably studied within the framework of Bartlett's theory.

REFERENCES


(Received November 12, 1973)