Retrieval-Induced Forgetting in Episodic Memory

Michael A. Ciranni and Arthur P. Shimamura
University of California, Berkeley

Participants learned the locations of 12 stimuli that were uniquely colored but could be grouped by shape (4 circles, 4 triangles, 4 crosses). Following the study, a retrieval-practice phase required participants to recall the colors of a subset of the stimuli (i.e., 2 circles, 2 triangles) using shape and location as cues. In a final test, participants recalled the colors of all 12 stimuli. Compared with the control set of stimuli (i.e., 4 crosses), memory was facilitated for practiced items but impaired for related items, which were not practiced but shared the same shape group. Across experiments, retrieval-induced forgetting was observed for different perceptual groupings and for different cuing procedures. The effect, however, required retrieval of information during the interpolated phase. Providing extra presentations did not disrupt memory for related items.

Interference effects can also be caused by strengthening memory for a subset of initially learned information prior to test. In studies of the list-strength effect, memory for relatively "weak" items is impaired by multiple presentations of "strong" items within the same list (Ratcliff, Clark, & Shiffrin, 1990). Also, in the verbal overshadowing paradigm, recognition memory for a face is actually impaired when participants are required to recollect the face by generating a verbal description of it before a test of face recognition (Dodson, Johnson, & Schooler, 1997; Schooler & Engstler-Schooler, 1990).

Similar effects can be produced by strengthening information at the time of testing. For example, the probability of recalling an item declines as a function of the number of responses already elicited, a phenomenon known as output interference (Rundus, 1973; Smith, 1971). Various theories (e.g., Raaijmakers & Shiffrin, 1981; Rundus, 1973) suggest that retrieval of an item during the test strengthens its representation. This newly strengthened item then interferes with the retrieval of subsequent study items. A related phenomenon is part-set cuing. Presenting a subset of previously learned items as retrieval "cues" often impairs recall for the remaining information (Roediger, 1973; Slamecka, 1968). Studies of the part-set cuing effect have shown that retrieval of both recently learned information (e.g., word lists) and remote memories (e.g., names of the 50 U.S. states) can be subject to this interference effect (J. Brown, 1968; Karchmer & Winograd, 1971; for a review, see Nickerson, 1984). Mueller and Watkins (1977) demonstrated that part-set cuing effects can be category specific. That is, participants in a cued-recall test were impaired at recalling the remaining information when given a partial set of exemplars from the same category but were not impaired when given a set of exemplars from a different category.

The retrieval-practice paradigm (M. C. Anderson, Bjork, & Bjork, 1994) has been shown to be a useful method to study interference effects. The paradigm is similar to the list-strength and part-list cuing paradigms except that the strengthening of partial information occurs during a separate interpolated phase rather than during the study or test phase.
phases. In this paradigm, participants first study a set of category–exemplar word pairs (e.g., Fruit–Orange, Fruit–Banana, Tool–Pliers). Following learning, a practice phase is presented in which participants are asked to retrieve only some of the exemplars from category–letter cues (e.g., Fruit–Or—__). After this retrieval-practice phase, there is a final test phase in which participants recall all of the exemplars with category–letter cues. It is interesting to note that recall is most impaired for items within the same category as those items presented during the retrieval-practice phase. For example, retrieving Orange when cued with Fruit–Or—— during the retrieval-practice phase impairs recall of the related item Banana when cued with Fruit–B—— more than the unrelated, or control, item Pliers when cued with Tool–P——.

The retrieval-practice paradigm offers several advantages over part-list cuing and other related interference paradigms. First, the retrieval-practice paradigm uses separate phases for study, interpolated practice, and test. In the list-strength studies and part-list cuing studies, the strengthening manipulations are confounded with initial learning or final testing. Another advantage is that the retrieval-practice procedure uses a within-subjects baseline to distinguish between generalized and item-specific interference effects. Although retrieval practice of some items is expected to impair memory in general for unpracticed items, the impairment is greater for the unpracticed related items than for the unpracticed control (unrelated) items. Moreover, the use of a cued-recall test controls for output order and thus helps to mitigate the effects of output interference. Without cued recall, participants would be expected to recall the practiced items first, thus producing an output interference effect on the remaining items independent of the retrieval-practice manipulation.

**Experiment 1**

In previous studies, the retrieval-practice paradigm has been used to assess interference effects for information that is already well organized in semantic memory. That is, the item-specific interference effects produced with the retrieval-practice paradigm may only occur for information that is well organized before the experiment begins. The purpose of the present investigation is to examine retrieval-induced forgetting for newly formed (episodic) associations. Rather than relying on preexisting category–exemplar relationships in semantic memory as a organizing principle, our experiments involved the acquisition of newly formed groupings based on perceptual features such as shape or color. Participants learned to associate each of 12 colored shapes (e.g., red circle) to a particular spatial location. In Experiment 1, each stimulus was assigned a unique color and location. Thus, perfect learning could be accomplished purely by using the association between color and location. Yet we also introduced a categorizing or grouping principle, such that four of the stimuli were circles, four were triangles, and four were crosses. After the learning of these stimulus–location associations, retrieval practice was given for a subset of the stimulus set (e.g., two of the circles, and two of the crosses). Finally, we tested stimulus–color memory for all 12 items.

To the extent that individuals use the shape dimension as a grouping principle, we predicted a retrieval-induced forget-ting effect in which the act of practicing a subset of items from a shape group (i.e., circles) would interfere with memory for nonpracticed items within the same shape group. Thus, recall for those items would be worse than for items in a shape group in which no members received retrieval practice (e.g., triangles).

**Method**

**Participants.** A total of 72 undergraduate students were tested and were paid $5 for their participation in the experiment. All participants reported normal or corrected-to-normal vision, and none reported any problems with color vision when asked before the experiment.

**Stimuli and design.** The stimuli were 12 colored shapes (four circles, four triangles, and four crosses), each displayed in a different color (blue, black, brown, gray, green, maroon, orange, pink, purple, red, teal, or yellow). The stimuli were positioned equally around a circular array with an internal diameter of 10° (see Figure 1). The location of each of the 12 stimuli was determined randomly at the beginning of the experiment, with the constraint that stimuli with the same shape would not be positioned in three consecutive locations around the array. The stimuli were displayed on a 14-in. monitor controlled by a Quadra 630 Macintosh using a program written in Hypercard 2.2.

The experiment was divided into three phases: a study phase, an interpolated practice phase, and a test phase. In the study phase, a continuous recognition paradigm was used in which participants learned the location of each stimulus. A stimulus was presented in the middle of the array, and participants attempted to learn the location of each stimulus. In the interpolated practice phase, four of the stimuli were given extra retrieval trials. Of the four practiced stimuli, two came from each of two shape sets (e.g., two circles, two triangles). Each stimulus was practiced three times by asking participants to identify the color of the stimulus, with feedback given immediately after each response. In the test phase, participants were asked to identify the colors of all stimuli. Thus, the test phase was essentially the same as the intervening practice phase except that every stimulus was tested once and no feedback was given.

The design enabled analysis of memory performance in three experimental conditions. The practiced condition involved the four stimuli that were used in the interpolated practice phase (e.g., the two circles and two triangles). The control condition involved the four stimuli that came from the shape set not used in the interpolated practice phase (e.g., the four crosses). The related condition involved the four stimuli that came from the shape sets used in the interpolated practice condition, but they themselves were not practiced (e.g., the remaining two triangles and remaining two circles).

To ensure that each combination of color and shape was used in each experimental condition, we counterbalanced which colors were assigned to which shape sets, which shape sets were practiced, and also which colors within a shape set were practiced. Thus, an individual colored shape (e.g., green circle) would ultimately be used in all three experimental conditions. This process produced 18 different versions of the interpolated practice phase.

**Procedure.** In the study phase, a trial began with a stimulus appearing in the center of the array. Participants were asked to
select a position at which the stimulus is located (see Figure 1). Participants were given 7 s to make a response. After the response, we provided feedback by showing the stimulus in its correct location. Also, an auditory tone indicated a correct or incorrect response. For the initial presentation of each stimulus, participants were essentially guessing the location. Thereafter, participants could use feedback to learn the position of each stimulus. A "dropout" procedure was used in which a stimulus was removed from the study phase after it was located successfully in seven trials. The order of presentation was randomly determined with the constraint that the same stimulus would not appear on consecutive trials, unless it was the last remaining stimulus to be learned. The study phase continued until participants had learned the correct location of each stimulus to a criterion of seven correct trials.

In the interpolated practice phase, participants were told that they would be given a test of their knowledge of the stimulus colors. Specifically, an uncolored stimulus was presented in its correct location, and participants were instructed to select the color that was associated with that stimulus (see Figure 1). Participants had 7 s to select a color by using a palette containing all 12 colors. Immediately following a response, we gave participants feedback by presenting the stimulus with its correct color. A tone was also presented to indicate a correct or incorrect response. Participants practiced retrieving the colors of four stimuli on each of three trials. The order of presentation was randomized. Participants were never informed that there was another test.

The actual test phase followed immediately after the retrieval-practice phase. The method used to assess memory on the test trials was essentially the same as that used in the retrieval-practice phase. On each test trial an uncolored stimulus appeared in its correct location, and participants selected the color associated with the stimulus. However, there was no feedback after the participant made a response (see Figure 1).
Results and Discussion

We analyzed the mean trials to learn the color–location associations in the study phase using single-variable analysis of variance (ANOVA). A significance level of .05 was used for all analyses. The mean trials to learn these associations for the practiced, related, and control items were 11.1, 11.0, and 11.3, respectively. The trials needed to learn the location of an item did not differ significantly, F(2, 142) = 1.76, MSE = 1.136. As the experimental manipulations do not occur until the interpolated practice phase, these findings assure equivalent initial learning of items in the three experimental conditions.

Table 1 displays the mean scores for percentage recall of practiced, related, and control items in the test phase. Average recall for the different item types was analyzed using single-variable ANOVA. There was a significant main effect of item type, F(2, 142) = 29.90, MSE = 0.575. Planned comparisons showed that memory was better for practiced items than for control items, F(1, 142) = 21.58, MSE = 0.550. More interestingly, memory for related items was significantly poorer than memory for control items, F(1, 142) = 8.90, MSE = 0.641.

Previous studies using semantically related associates have shown that retrieval practice impairs memory for semantically related items whose associations were formed before the study phase (M. C. Anderson et al., 1994). In this study, items were related by newly learned associations among shapes, colors, and locations. Thus, the interfering effects of retrieval practice were extended to associations that were acquired in a single learning episode.

Experiment 2

The purpose of this experiment was first to replicate the results of the previous experiment and to extend our results by demonstrating retrieval-induced forgetting for a different perceptual grouping. One condition was identical to Experiment 1, in that the stimuli could be grouped by shape but had a unique color. The stimuli in the second set could be grouped by color, but each stimulus had a unique shape. We expected to find similar patterns of retrieval-induced forgetting for items grouped by color and items grouped by shape.

Method

Participants. We tested 108 undergraduates who were paid $5 for their participation in the experiment. All participants reported normal or corrected-to-normal vision, and none reported any problems with color vision when asked before the experiment.

Stimuli and design. Two stimulus sets were used. In the first set, we used the colored shapes that were used in Experiment 1. That is, four circles, four triangles, and four crosses, each with a different color, were assigned positions in a circular array. Thus, the grouping feature for stimuli in this set was shape. In the second set, the grouping feature was color. Four green, four orange, and four purple stimuli were assigned a unique shape (circle, crescent, cross, diamond, drop, hexagon, hourglass, leaf, oval, star, trapezoid, or triangle). Each participant was presented with only one of the stimulus sets, and the participants were evenly divided between these two grouping-feature conditions.

The experiment used the same three phases as Experiment 1: study phase, interpolated practice phase, and test phase. The design used in the shape grouping condition was exactly the same as that used in Experiment 1. An analogous design was used in the color grouping condition. Thus, for the interpolated practice phase, items were presented from each of two color sets (e.g., two orange shapes, two green shapes). Each of these stimuli was practiced three times by asking participants to identify the shape of the stimulus (with feedback) when cued by the color and location. The test phase again required participants to identify the shapes of all of the stimuli, and no feedback was given.

We analyzed memory performance with respect to the two different grouping-feature conditions (shape or color) and with respect to the three experimental conditions used in Experiment 1: practiced items, related items, and control items. As in Experiment 1, the practiced condition referred to those stimuli that were presented in the interpolated practice phase. The related condition referred to those unpracticed stimuli that shared the same grouping feature as the practiced stimuli. The control condition referred to those unpracticed stimuli that did not share the same grouping feature as the practiced stimuli.

We counterbalanced the combinations of color and shape for both stimulus sets to ensure that each combination was used in each experimental condition. The counterbalancing was the same as described in Experiment 1 for the shape grouping condition. For the color grouping condition, we counterbalanced which shapes were assigned to which color sets, which color sets were practiced, and also which shapes within a color set were practiced. This process again produced 18 different versions of the interpolated practice phase.

Table 1

Percentage Recall of Practiced, Related, and Control Items Across Experiments 1–5

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Grouping feature</th>
<th>Type of study</th>
<th>Type of practice</th>
<th>Type of test</th>
<th>Practiced</th>
<th>Related</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp. 1</td>
<td>Shape</td>
<td>Learn location</td>
<td>Retrieve color</td>
<td>Color</td>
<td>95</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>Exp. 2</td>
<td>Shape</td>
<td>Learn location</td>
<td>Retrieve color</td>
<td>Color</td>
<td>88</td>
<td>71</td>
<td>76</td>
</tr>
<tr>
<td>Exp. 3</td>
<td>Color</td>
<td>Learn location</td>
<td>Retrieve shape</td>
<td>Shape</td>
<td>93</td>
<td>72</td>
<td>80</td>
</tr>
<tr>
<td>Exp. 4</td>
<td>Color</td>
<td>Learn location</td>
<td>Retrieve shape</td>
<td>Location</td>
<td>94</td>
<td>86</td>
<td>91</td>
</tr>
<tr>
<td>Exp. 5</td>
<td>Color</td>
<td>Learn location</td>
<td>Retrieve location</td>
<td>Location</td>
<td>96</td>
<td>83</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Color</td>
<td>Learn location</td>
<td>Present stimulus</td>
<td>Shape</td>
<td>90</td>
<td>67</td>
<td>77</td>
</tr>
</tbody>
</table>

Note. Exp. = experiment.
Procedure. For participants in the shape grouping condition, the experiment was conducted exactly as it was in Experiment 1. The procedure for participants in the color grouping condition was the same as Experiment 1 except for the following changes. In each retrieval-practice trial, a colored question mark was presented in a particular location. The color of the question mark matched the correct color of the shape assigned to that location. Participants were instructed to select the shape that was presented at that location during the study phase. Responses were made by selecting a shape from a palette containing all 12 shapes. Following each response, we gave participants feedback by presenting the correct colored shape at that location and playing a tone indicating whether their shape choice was correct or incorrect.

For both grouping conditions, the procedures used in this interpolated practice phase were also used in the test phase. However, for the test phase no feedback was given. After a response was made, the computer initiated the next trial without revealing the correct answer.

Results and Discussion

We analyzed the effects of grouping feature and item type on mean trials to learn the associations in the study phase using a 2 x 3 ANOVA. A significance level of .05 was adopted for all analyses. For stimuli grouped by shape, the mean trials to learn the color–location associations of practiced, related, and control items were 11.5, 11.3, and 11.0, respectively. For stimuli grouped by color, the mean trials to learn the shape–location associations of practiced, related, and control items were 10.6, 10.4, and 10.3, respectively. There was no significant effect of item type, F(2, 212) = 2.54, MSE = 1.270. However, stimuli grouped by shape required significantly more trials to learn than stimuli grouped by color, F(1, 106) = 9.70, MSE = 5.819. The dropout procedure used in the study phase should have minimized these differences in the initial learning, however. This assertion is supported by the finding that the percentage of practice items recalled correctly in the interpolated practice phase was not significantly different between the shape and color grouping feature conditions (91% vs. 88%), F(1, 106) = 0.91, MSE = 138.185. There was no interaction between grouping feature and item type, F(2, 212) = 0.20, MSE = 1.270.

The mean percentage recall for practiced, related, and control items in both grouping feature conditions is presented in Table 1. The effects of grouping feature and item type on recall were analyzed using a 2 x 3 ANOVA. There was a main effect of item type, F(2, 212) = 32.74, MSE = 0.620. Two planned comparisons were conducted to examine the nature of this effect. The first comparison showed that memory for practiced items was significantly better than memory for control items, F(1, 106) = 33.69, MSE = 0.401. Practice on items grouped by either feature improved performance on the final test. The second planned comparison showed that memory for related items was significantly worse than memory for control items, F(1, 106) = 6.50, MSE = 0.559. There was no main effect of grouping feature, F(1, 106) = 0.95, MSE = 1.300, nor was there an interaction between item type and grouping feature, F(2, 212) = 0.33, MSE = 0.491.

In summary, the retrieval-induced forgetting effect was reliable regardless of whether the items were grouped by shape or grouped by color. It is unclear, however, whether this effect is specific to the manner in which the associations are tested. In Experiments 1 and 2, participants learned the items in the study phase by using color and shape cues to retrieve the locations. In contrast, the interpolated practice phase used location cues combined with shape or color to retrieve the remaining feature (color or shape). The test phase used the same procedure as the interpolated practice phase.

For the next two experiments, we used two different test conditions. In one condition, the test conditions were similar to the interpolated practice conditions, as with Experiments 1 and 2. In the other condition, the test conditions were more similar to the study conditions. This manipulation allowed us to address two important concerns.

The first concern is that the interpolated practice effects may occur because the test conditions match the conditions at practice better than the conditions at study. The encoding specificity principle (Tulving & Thomson, 1973) predicts that memory will be best when conditions during retrieval are most similar to the conditions during encoding. If this is the case, then changing the final test to match the study conditions may negate this interpolated practice effect.

The second concern involves the extent of the interpolated practice effect. It is possible that interpolated practice affects only the retrieved feature of the practiced and related stimuli. That is, retrieval of shape information in the interpolated practice phase facilitates memory for the shapes of the practiced stimuli and impairs memory for the shapes of the related stimuli, but it may not affect memory for the other features of these stimuli (i.e., their colors and locations). On the other hand, interpolated practice may affect memory for other features of the practiced and related stimuli. That is, retrieving the shape of a stimulus may cause memory for the shapes of related stimuli to be impaired, but it may also cause memory for other features such as location to be impaired in these related stimuli as well. If this is the case, then interpolated practice should produce the same effects when the tested feature is different from the practiced feature.

Experiment 3

In this experiment, we used the same study and interpolated practice conditions that we used for the stimuli grouped by shape in Experiment 2. However, there were two different test conditions. In one condition, the test required participants to retrieve the shape of a stimulus, as they had done in the practice phase. In the other condition, the test required participants to retrieve the location of a stimulus, as they had done in the study phase. We predicted that if retrieval-induced forgetting is not dependent on a match between practice and test conditions, then the effect would occur in both test conditions.

Method

Participants. We tested 108 undergraduate students who fulfilled a psychology class requirement for their participation in the
experiment. All participants reported normal or corrected-to-normal vision, and none reported any problems with color vision when asked before the experiment.

**Stimuli and design.** The stimuli were grouped by color (i.e., four green, four orange, and four purple), and each had a different shape. The stimuli were positioned in the same circular array as in Experiments 1 and 2. The color–shape combinations for the stimuli were counterbalanced across the experimental conditions using the same procedures described for the color grouping condition of Experiment 2.

There were two test conditions in this experiment: a shape retrieval test condition and a location retrieval test condition. In the shape retrieval test condition, the design was exactly the same as the color grouping condition used in Experiment 2. That is, participants retrieved shape information when cued by location. In the location retrieval test condition, the location of the array, and the participant selected the location for that stimulus. That is, participants retrieved location information when cued by shape. Thus, each test trial in this latter condition was similar to a trial in the study phase, except no feedback was given. The study and practice phases were the same between the two test conditions. Each participant performed only one of the two test conditions, and the participants were evenly divided between these two test conditions.

**Procedure.** The procedure for the shape retrieval test condition was identical to the procedures used for the stimuli grouped by color in Experiment 2. The procedure for the location retrieval test condition had the following changes: On each test trial, 1 of the 12 stimuli would appear in the center of the array in its correct color. The participant was instructed to select the position that he or she remembered to be the location of that stimulus. Once the participant had made a selection, a new trial began. Each stimulus was presented exactly once in the test phase.

**Results**

We analyzed the effects of test condition and item type on mean trials to learn the shape–location associations in the study phase using a $2 \times 3$ ANOVA. A significance level of .05 was used for all analyses. For the shape retrieval test condition, the mean trials to learn the practiced, related, and control items were 10.9, 10.9, and 10.9, respectively. For the location retrieval test condition, the mean trials were 10.8, 10.8, and 10.7, respectively. There was no difference in the trials needed to learn the locations of items between the two test conditions, $F(1, 106) = 0.01$, $MSE = 7.936$. In addition, the effect of item type was not significant, $F(2, 212) = 0.23$, $MSE = 0.971$, nor was there an interaction between test condition and item type, $F(2, 212) = 1.53$, $MSE = 0.971$. In the practice phase, participants in the shape and location retrieval test conditions recalled 87% and 84% of practice items, which was not significantly different, $F(1, 106) = 0.91$, $MSE = 204.754$. These findings were expected because memory for these items was not manipulated until after the study phase.

The mean recall scores for each of the practice and type of test conditions are presented in Table 1. We analyzed the effects of test condition and item type using a $2 \times 3$ ANOVA. As in Experiments 1 and 2, there was a significant effect of item type, $F(2, 212) = 22.77$, $MSE = 0.405$. Two planned comparisons showed that memory for practiced items was significantly better than memory for control items, $F(1, 106) = 19.16$, $MSE = 0.296$, but memory for related items was significantly worse than memory for control items, $F(1, 106) = 6.85$, $MSE = 0.530$.

The impairment in memory for related items holds when the test conditions are analyzed separately. In the location retrieval test condition, memory for related items was significantly worse than memory for control items, $F(1, 53) = 4.78$, $MSE = 0.234$. In the shape retrieval test condition, memory for related items was marginally worse than memory for control items, $F(1, 53) = 3.24$, $p = .078$, $MSE = 0.827$, due to the increased error variance in this condition offsetting the larger observed impairment. There was no interaction between test condition and memory for related and control items, $F(1, 106) = 0.31$, $MSE = 0.530$.

Memory for all items in the location retrieval test condition was significantly better than memory for all items in the shape retrieval test condition, $F(1, 106) = 19.27$, $MSE = 0.925$. In addition, there was also an overall interaction between test condition and item type, $F(2, 212) = 5.36$, $MSE = 0.405$. Participants in the shape retrieval test condition showed greater memory for practiced items than control items, $F(1, 53) = 19.01$, $MSE = 0.410$, but participants in the location retrieval test condition did not, $F(1, 53) = 1.83$, $MSE = 0.182$. The amount of facilitation for practiced items did differ significantly in comparing memory for practiced and control items between test conditions, $F(1, 106) = 8.27$, $MSE = 0.296$.

**Experiment 4**

For this experiment, we used the same test conditions used in Experiment 3 but changed the study and interpolated practice phases. The study phase involved learning the shapes of all the stimuli, and the interpolated practice phase involved retrieving the location of items. These two phases are the reverse of what was used in Experiment 3. The predictions are the same as for Experiment 3: The retrieval-induced forgetting effect should be observed in both test conditions.

**Method**

**Participants.** We tested 108 undergraduate students who fulfilled a psychology class requirement for their participation in the experiment. All participants reported normal or corrected-to-normal vision, and none reported any problems with color vision when asked before the experiment.

**Stimuli and design.** The stimuli were grouped by color (i.e., four green, four orange, and four purple), and each had a different shape. The stimuli were positioned in the same circular array as in Experiments 1–3. The color–shape combinations for the stimuli were counterbalanced across the experimental conditions using the same procedures described for the color grouping condition of Experiment 2.

There were two test conditions in this experiment: a shape retrieval test condition and a location retrieval test condition. These conditions were the same as the test conditions used in Experiment 3. The study and practice phases were the same between the two test conditions. Participants were evenly divided between the test conditions.
Procedure. On each trial of the study phase, a colored question mark appeared in 1 of 12 locations. The color of the question mark matched the color of the stimulus assigned to that location. Participants were instructed to select a shape that they thought might belong at that location, using a palette containing all 12 shapes. Participants received feedback after each response by presenting the correct shape at that location and playing a tone indicating whether their shape choice was correct or incorrect. As before, a dropout procedure was used in which a stimulus was removed from the study phase after its shape was chosen successfully in seven trials. The order of presentation was randomly determined with the constraint that the same stimulus would not appear on consecutive trials unless it was the last remaining stimulus to be learned. The study phase continued until participants had learned the correct shape at each location to a criterion of seven correct trials.

The procedure for the interpolated practice phase was similar to the procedure for the study phase used in Experiments 1–3. A colored shape appeared in the center of the screen, and the participant clicked on the location where the stimulus had been seen in the study phase. After making a choice, the participant was shown the correct location of the stimulus. As with the previous experiments, only 4 of the 12 stimuli (2 stimuli from two different color sets) were shown in the interpolated practice phase, and each of these items was presented three times during the interpolated practice phase. The procedures for both the location and shape retrieval test conditions were the same as those used for Experiment 3.

Results and Discussion

We analyzed the effects of test condition and item type on the mean trials to learn the location–shape associations during the study phase using a 2 × 3 ANOVA. A significance level of .05 was adopted for all analyses. The mean trials to learn the practiced, related, and control items were, respectively, 11.7, 11.8, and 11.6 for the shape retrieval test condition and 11.6, 11.7, and 11.6 for the location retrieval test condition. The number of trials needed to learn the shapes of these items did not vary significantly between the two test conditions, $F(1, 106) = 0.08, MSE = 12.216$. In addition, there was no effect of item type, $F(2, 212) = 0.63, MSE = 1.776$, and there was no interaction between test condition and item type, $F(2, 212) = 0.15, MSE = 1.776$. Participants in both the shape and location retrieval test conditions recalled 88% of the practice items in the interpolated practice phase. As with Experiment 3, we did not expect to find any differences here because memory for these items was not manipulated until after the study phase.

The mean recall scores for practiced, related, and control items in each test condition are presented in Table 1. The effects of test condition and item type were analyzed using a 2 × 3 ANOVA. Once again, there was a significant effect of item type, $F(2, 212) = 24.29, MSE = 0.278$. Two planned comparisons showed that memory for practiced items was significantly better than memory for control items, $F(1, 106) = 13.64, p < .001, MSE = 0.266$, but memory for related items was significantly worse than memory for control items, $F(1, 106) = 14.61, MSE = 0.214$. There was no main effect of test condition, $F(1, 106) = 0.17, MSE = 0.643$, nor was there an interaction between test condition and item type, $F(2, 212) = 0.40, MSE = 0.278$.

When the test conditions are analyzed separately, memory for related items was still significantly worse than memory for control items in both the location retrieval test condition, $F(1, 53) = 4.87, MSE = 0.190$, and in the shape retrieval test condition, $F(1, 53) = 9.95, MSE = 0.238$. There was no interaction between test condition and memory for related and control items, $F(1, 106) = 0.78, MSE = 0.214$.

Taken together, the results of Experiments 3 and 4 suggest that the retrieval-induced forgetting effect does not depend on any similarity between the practice and test conditions. The location retrieval test itself may produce a slightly smaller retrieval-induced forgetting effect than the shape retrieval test (5% vs. 8%). However, this difference was apparent both when the location retrieval test matched the practice conditions and when the location retrieval test matched the study conditions. Thus, this disparity cannot be explained in terms of a match between the practice and test conditions. The location retrieval test may simply be an easier test than the shape retrieval test.

Experiment 5

In this experiment, we investigated whether the interference effect is specifically caused by retrieval of related information during the interpolated practice phase or whether any form of strengthening of memory between study and test will cause interference. To test this possibility, we used two different interpolated practice conditions. In the retrieval-practice condition, participants were asked to retrieve the shape of an item when cued with its location. This condition is analogous to the conditions of Experiments 2 and 3. In the extra-presentations condition, participants were shown an item in its correct location and were told to focus their attention on the shape of the item. We hypothesized that if retrieval is necessary to produce the memory impairment for related items, then memory for related items will be impaired in the retrieval condition but not in the extra-presentations condition.

Method

Participants. We tested 108 undergraduate students who fulfilled a psychology class requirement for their participation in the experiment. All participants reported normal or corrected-to-normal vision, and none reported any problems with color vision when asked before the experiment.

Stimuli and design. The 12 stimuli used were grouped by color (i.e., four green, four orange, and four purple), and each stimulus had a different shape. The stimuli were positioned in the same circular array as in Experiments 1–4. The color–shape combinations for the stimuli were counterbalanced across the experimental conditions using the same procedures described for the color grouping condition of Experiment 2.

There were two interpolated practice conditions in this experiment. In the retrieval-practice condition, the design was the same as the color grouping condition used in Experiment 2. That is, a colored question mark would appear in a particular location, and participants selected the shape presented at that location in the study phase. In the extra-presentations condition, the interpolated practice phase was changed so that on each practice trial a colored shape appeared in its correct location, and participants were
instructed to focus their attention on the shape of the stimulus. Thus, no overt retrieval of information was required in this condition, although presumably the focusing of attention on the stimulus would strengthen the participant's memory for that stimulus. The study and test phases were the same between the two practice conditions. Each participant was assigned to only one of the two interpolated practice conditions, and the participants were evenly divided between these conditions.

Procedure. In the retrieval-practice condition, the experiment was run the same as the color grouping condition of Experiment 2. On each trial of the extra-presentations condition, a stimulus would appear in its correct location for 3 s. Participants were instructed to focus their attention on the shape of the stimulus and to try mentally associating the shape of the stimulus with its color and its location. Because a retrieval-practice trial would produce greater strengthening than an extra-presentations trial (Jacoby, 1978; Slamecka & Graf, 1978), we presented each stimulus in the extra-presentations condition four times to try to equate the amount of strengthening between the two practice conditions. The stimuli were presented in a random order. There was an intertrial interval of 1 s.

Results and Discussion

The effects of practice condition and item type on the time to learn the locations of the stimuli were analyzed using a 2 X 3 ANOVA. A significance level of .05 was used for all analyses. The mean trials to learn the practiced, related, and control items from each stimulus set were, respectively, 10.9, 10.9, and 10.9 for the retrieval-practice condition and 10.6, 10.5, and 10.8 for the extra-presentations condition. There was no difference between the two practice conditions, F(1, 106) = 0.65, MSE = 9.965. In addition, there was no effect of item type, F(2, 212) = 0.55, MSE = 1.137, nor was there an interaction between practice condition and item type, F(2, 212) = 0.84, MSE = 1.137. As expected, memory for the practiced, related, and control items did not differ significantly because there were no experimental manipulations at this point.

The mean percentage recall for each of the practice and grouping conditions is presented in Table 1. The effects of practice condition and item type on recall were analyzed using a 2 X 3 ANOVA. As before, there was a main effect of practice condition, F(1, 106) = 0.95, MSE = 1.157, nor was there an overall interaction between practice condition and item type, F(2, 212) = 0.16, MSE = 0.491. This interaction, however, may have been diminished because we did not expect differences in memory performance for the practiced and control items across the two types of practice. Only memory for related items should be affected by the practice manipulation.

To assess the effect of practice condition on memory performance, we restricted our analyses to the related and control conditions. There was a significant interaction in memory for related and control items and practice condition, F(1, 106) = 4.17, MSE = 0.537. Two planned comparisons for determining the nature of this interaction showed that in the retrieval-practice condition, correct recall of related items was significantly lower than for control items, F(1, 53) = 8.86, MSE = 0.461. However, there was no difference in recall for related and control items for the extra-presentations condition, F(1, 53) = 0.02, MSE = 0.613. Thus, practice through retrieval of learned information disrupts memory for related items compared with control items, whereas practice through additional study of learned information does not disrupt memory for related items.

We also analyzed memory performance for practiced items to see if the two types of practice produced differences in the amount of strengthening. Memory for practiced items was significantly better than memory for control items, F(1, 106) = 27.82, MSE = 0.660. There was no difference in memory for practiced and control items between the two practice conditions, F(1, 106) = 0.31, MSE = 0.731, and no interaction in the difference in memory performance between the practiced and control conditions and the type of practice, F(1, 106) = 0.18, MSE = 0.660. Thus, both types of practice seem to improve memory for practiced items over control items, and there was no difference in the amount of improvement.

In summary, memory for the practiced stimuli was greater than memory for the control stimuli, regardless of how this practice was implemented. The type of practice did, however, have an effect on the related stimuli. Retrieval practice of practiced items caused memory for related items to be poorer than memory for the control items, whereas extra presentations of practiced items did not impair memory for the related items compared with the control items.

General Discussion

The findings from these five experiments extend the domain of retrieval-induced interference effects to recently learned, episodic information. Experiments 1 and 2 demonstrated this basic effect using stimuli that could be grouped perceptually by color or by shape. Retrieving information about items in one perceptual group impaired subsequent recall for the remaining members of that group. These findings suggest that participants adopt an organizational strategy in which stimuli are grouped by common perceptual features, even when this grouping was irrelevant. That is, in Experiment 1 each stimulus had a unique color and location, and thus performance could be completely successful without using shape information. Nonetheless, the seemingly redundant grouping of stimuli by shape was incorporated such that it led to memory interference rather than facilitation.

Experiments 3 and 4 analyzed the retrieval-induced forgetting effect by manipulating the cues used at test. In one condition, test cues matched the cues used in the interpolated practice phase. In the other condition, test cues matched the cues used during the initial learning. This latter condition is also important because the feature retrieved during the final test (i.e., shape or location) was not the same as the feature retrieved during the interpolated practice. Despite changes in the manner in which items were assessed, both test conditions produce comparable degrees of interference. The size of the effect seemed to depend more on the types of cues used (i.e., location or shape cues) rather than the similarity of these cues to those used in previous phases of the experiment.
These results demonstrate two additional properties. First, the effect does not depend on a match between practice and test conditions. Second, the interference produced by interpolated practice is not limited to a single feature. Even when the feature retrieved during the final test is different from the feature retrieved during the interpolated practice, memory for related items is still affected. Because the test cues are different from the practice cues, the cues should not elicit retrieval of the practiced items. Thus, it is unlikely that this impairment is directly related to greater activation of the practiced items, as would be predicted by blocking theories of retrieval interference. Rather, the interference effect may instead be caused by a degradation or suppression of the related items.

Experiment 5 demonstrated that retrieval-induced forgetting depends critically on the act of retrieving information during the interpolated practice phase. The mere exposure of study items during an interpolated phase did not cause interference, even though the condition facilitated memory performance for practiced items. However, only the group given retrieval practice exhibited impaired memory for related items. This finding is particularly interesting, in that it suggests that there may be qualitative differences between merely presenting information and actively retrieving it. Such differences are not generally accounted for in many current models of memory that use a strength-based, ratio-rule system for information retrieval (J. R. Anderson, 1983; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1981; Rundus, 1973), which do not consider how information is strengthened in memory.

For example, these results at first glance appear to be inconsistent with the list-strength effect (Ratcliff et al., 1990). In a typical list-strength paradigm, participants are given lists of words containing “strong” items presented for a long duration or “weak” items presented for a short duration. These lists can contain only strong items, only weak items, or a mixture of strong and weak items. When a participant is given a mixed list of words containing strong and weak items, recall of the weak items is proportionately worse on the mixed list compared with a pure list containing only weak items. This effect is predicted by the search of associated memory (SAM) model and other ratio-rule models of retrieval. The strong items in a mixed list have a proportionately greater chance of being retrieved than the weak items on any given retrieval attempt; thus, fewer weak items will be recalled from a mixed list than from a pure list. If the stimulus color groupings are analogous to the lists, the extra-presentations condition could be producing mixed lists of strong practiced items and weak related items. Thus, one might predict poorer performance for the weak related items compared with the pure group of weak control items.

However, other studies have suggested that the list-strength effect itself is actually an artifact of output interference. If a free-recall test is used, strong items are presumably recalled first, and it is the retrieval of these items during the test that is suppressing the weak items. However, if the output order of the items is controlled by using a cued-recall test, these order biases can be removed. Bäuml (1997) showed that when the output order is controlled, the list-strength effect disappears. Because the present set of experiments also used a cued-recall test, it is not surprising that we failed to find a list-strength effect.

Taken together, the present research and previous studies of the list-strength effect and output interference suggest that it is the act of retrieval that is responsible for these effects. M. C. Anderson et al. (1994; M. C. Anderson & Spellman, 1995) suggested that retrieval of information involves the suppression of potential competitors. For example, a semantic cue (e.g., Fruit—Or——) is presumed to activate other, irrelevant semantic associates that must be inhibited. This inhibition effect would occur directly as a result of attempting to retrieve information when potential competitors are active.

A recent study by M. C. Anderson, Bjork, and Bjork (in press) provides evidence for this hypothesis. Participants in the retrieval-practice phase engage in either competitive retrieval using category-stem cues to retrieve an exemplar (e.g., Fruit—Or——), or noncompetitive retrieval using exemplar-stem cues to retrieve a category (e.g., Fr———Orange). Only the competitive condition produced memory impairments for the unpracticed but related exemplars. Because there are presumed to be no competing categories for an exemplar, the noncompetitive condition should not produce any inhibition. These results are similar to what we found in the extra-presentations condition of Experiment 5. That is, the mere presentation of items did not involve retrieval. Thus, there is no need to inhibit any competing items, and there would be no impairment in memory for related items.

The deleterious effects of semantic competitors have been well documented in semantic retrieval tasks. Blaxton and Neely (1983) gave participants categories plus single-letter stems (to isolate a particular exemplar from the category) and asked participants to generate an exemplar of that category beginning with the letter. Before recalling these targets, however, participants were also asked to either read or generate other exemplars from the same or a different category. Blaxton and Neely showed that the reaction times to generating the target exemplar were faster if the participant first read other exemplars from the same category compared with reading exemplars from a different category. However, the reaction times were slower if the participants had first generated other exemplars from the same category. This effect is similar to the results of Experiment 5 in that interpolated practice produced interference only when participants were required to retrieve the practiced exemplars. This effect is also somewhat similar to the output interference effects described previously, but in this case the interference effect is acting on retrieval of purely semantic information.

Other semantic retrieval tasks have demonstrated how semantically related primes can produce interference. A. S. Brown (1979) gave participants definitions of words (e.g., “to swallow up or eat greedily”) and asked the participants to say the word matching the definition (e.g., gobble). Following the definition, participants were also exposed to a prime word that was semantically related (e.g., cram) or unrelated (e.g., feud). Several studies have shown that
participants were slow in providing the definition when the prime was semantically related (Bowles & Poon, 1985; A. S. Brown, 1979; but see Roediger, Neely, & Blaxton, 1983). Similar results were reported by La Heij, Starreveld, and Steehouwer (1993) for participants who were given a general knowledge question (e.g., the color of tomatoes) and asked to retrieve the missing semantic information. Participants were slower to answer the question when it was followed by a semantically related prime (e.g., *yellow*) than an unrelated prime (e.g., *house*).

More basic tasks requiring semantic knowledge can also be impaired by semantically related primes. In picture-naming tasks, participants are slower to name the picture if it is accompanied by a semantically related prime than by an unrelated prime (Glaser & Dungelhoff, 1984; La Heij, 1988; Lupker, 1979; Starreveld & La Heij, 1996; Underwood, 1976; Wheeldon & Monsell, 1994). Semantic primes can also produce interference in lexical decision tasks. Typically, semantic primes produce facilitation (e.g., Meyer & Schvaneveldt, 1971; Neely, 1977; for reviews, see Carr, 1986; Ratcliff & McKoon, 1988). However, if the experiment is designed so the participant must retrieve semantic information about the prime (e.g., recall the definition of a prime), then a semantically related prime will slow participants on the lexical decision task (Dagenbach, Carr, & Barnhardt, 1990; Dagenbach, Carr, & Wilhelmson, 1989). Taken together, the evidence from these paradigms suggests that semantic competitors can act to interfere with the retrieval of other semantic information.

The retrieval-practice paradigm developed by M. C. Anderson and colleagues is an interesting case, because the study items have both episodic and semantic representations. The category–exemplar relationships used in the retrieval-practice procedure are “learned” in the study episode, but these relationships are also well–established, semantic associations that existed prior to the study phase of the experiment. Thus, the items used in this experiment essentially have two kinds of representations: (a) the episodic, contextual associations between category and exemplar produced in the study phase of the experiment (i.e., *Fruit–Banana* is a word pair seen in this experiment) and (b) the well-established semantic associations of each exemplar to the category (i.e., *fruit* is a *banana*).

It is unclear whether the inhibition effects produced by retrieval practice are acting on the semantic representations or episodic, contextual representations. However, there is some evidence that the effect may be operating on a semantic level. The second-order inhibition experiments reported in Experiment 2 of M. C. Anderson and Spellman (1995) suggest that semantic information that was not even part of the initial learning phase can be affected. For example, participants in one study were given *Green–Emerald*, *Green–Lettuce*, *Soup–Mushroom*, and *Soup–Minestrone* to study. The participants then retrieved *Green–Emerald* in the retrieval-practice phase. This manipulation caused memory for *Green–Lettuce* to decrease, relative to *Soup–Minestrone*, but also caused memory for *Soup–Mushroom* to decrease relative to *Soup–Minestrone*. According to Anderson and Spellman, retrieval practice of *Green–Emerald* causes the representation for *Lettuce* to be suppressed. In addition, lettuce and mushroom are both vegetables (in the minds of typical college undergraduates, at least), although this semantic association is not an explicit part of the experimental context. Thus, because part of the *Lettuce* representation is the semantic information that lettuce is a vegetable, this piece of semantic information (vegetable) is inhibited in *Mushroom* as well, producing a memory impairment for *Soup–Mushroom*. For these second-order retrieval-induced forgetting effects to be produced, the inhibition may be occurring at this semantic level.

The present set of experiments extends this work by using information that is perceptually related rather than semantically related. Whereas the associations between categories and exemplars are already well established, the associations between a particular color and shape are completely arbitrary. Thus, the network of associations comprising the stimulus sets in these experiments could only be acquired during the study phase. Nonetheless, the representations established in the study phase were sufficient to produce retrieval-induced forgetting.

These findings of retrieval-induced forgetting with episodic information shed new light on the mechanisms involved in more traditional retroactive interference paradigms, such as A–B, A–C paired-associate learning. In both the present study and in traditional retroactive interference studies, the effects of interpolated learning were assessed for newly learned associations, as opposed to previously learned semantic knowledge. In many of these studies, the interpolated A–C learning involved a series of cued-recall trials until the participants had learned the A–C list to a certain criterion. These cued-recall trials are analogous to the interpolated learning used in the present study and suggest that the learning of C associations to the cue A may involve the inhibition of the original B associations. Indeed, the response-set suppression hypothesis put forth by Postman et al. (1968) to explain many of the results on tests of paired-associate learning is quite similar to the inhibitory explanation proposed by M. C. Anderson et al. (1994), although the importance of retrieval processes in producing this suppression was not a component of the response-set suppression hypothesis advocated by Postman et al.

Retrieval-induced forgetting may also be used to explain other memory phenomena, such as the verbal overshadowing effect discovered by Schooler and Engstler-Schooler (1990). They found that the act of making a verbal description of a previously presented face impaired subsequent recognition of that face. It is possible that while generating the verbal description, people retrieve various aspects of the face (e.g., color of hair and eyes, expression of mouth) and in doing so suppress other aspects of the face (e.g., placement of cheekbones, hairline and style of hair, and shape of nose). The suppression of these features may thus impair subsequent recognition of the face. Consistent with this view, Dodson et al. (1997) found that when participants were given the task of describing a familiar face, such as a parent’s during the interpolated phase, this task was sufficient to impair face recognition. That is, the act of retrieving
aspects of any face appeared to inhibit features of the target face.

It is easy to imagine how such distortions in face memory might have disturbing implications for eyewitness testimony. Shaw, Bjork, and Handal (1995) demonstrated how retrieval-induced forgetting can occur in other aspects of eyewitness testimony to corrupt eyewitness memory. In their study, participants viewed a series of slides containing numerous objects in a dorm room. Following the viewing, there was an “interrogation phase” in which half of the participants were then asked a series of written questions about some of the objects they had seen in the slides, and the other half of the participants were given some problems to solve. After the interrogation phase, there was a final test in which both groups were asked to recall as many items in the room as possible. The “interrogated group” remembered more items that they had been questioned about than the control group but remembered fewer items that they were not asked about. Thus, retrieval-induced forgetting may add an additional explanation for how inaccuracies can occur with eyewitness testimony. Note that unlike the misinformation studies (Loftus, 1975, 1979; McCloskey & Zaragoza, 1985), no misleading or erroneous information is involved; nonetheless, the eyewitness’s memory is still corrupted by the experimental manipulation.

In conclusion, retrieval-induced forgetting has been shown to occur for newly learned, episodic information just as it has been shown to occur for semantically related information. This finding suggests that organizational strategies involved in grouping items by common perceptual features can, under certain circumstances, interfere with subsequent memory rather than improve it. It is interesting to note that the interference effect was dependent on active retrieval of partial information during the retrieval-practice phase. When interpolated learning involved the mere presentation of partial information, it did not result in interference. These findings, taken together with those involving semantically related material, suggest that retrieval-induced forgetting is a general memory phenomenon that may be useful for advancing many contemporary theoretical and practical issues related to memory.

References
Melton, A. W., & Irwin, J. M. (1940). The psychology of human performance


Received September 30, 1997
Revision received May 11, 1999
Accepted May 12, 1999