

Processing strategies and the generation effect: Implications for making a better reader

PATRICIA ANN DEWINSTANLEY
Oberlin College, Oberlin, Ohio

and

ELIZABETH LIGON BJORK
University of California, Los Angeles, California

When presented with items that must be generated versus read at encoding, individuals typically remember better those items that they generated versus those that they only read. We examined whether—given the opportunity to experience such differential memorial consequences of generating versus reading—participants might change how they processed future to-be-read information. In a first set of two experiments, participants were able to profit from such an experience to the extent that a generation advantage was eliminated on subsequent memory tests of generated and read items. Two additional experiments demonstrated the critical nature of this experience in leading to improved processing of future to-be-read information and elimination of a generation advantage. We believe that these results relate to the characterization of the learner emerging from recent metacognitive research and have possible implications for how learners might be induced to process information more effectively.

The generation effect, which refers to the finding that we typically remember information better when we have taken an active part in producing it, rather than having it provided to us by an external source—as would be the case, for example, if we had to generate the word *kitten* from a word fragment (e.g., “k-tt-n”) versus being given the intact printed word to read—has been recognized in memory research for over 20 years (Jacoby, 1978; Slamecka & Graf, 1978). Moreover, since that time, the generation effect has proven to be quite robust and general, extending, for example, to mathematical problems (Lawson & Chinnappan, 1994; McNamara & Healy, 1995a, 1995b, 2000), answers to trivia questions (deWinstanley, 1995; Pesta, Sanders, & Murphy, 1999; Peynircioğlu & Mungan, 1993), and reading comprehension (Wittrock, 1990).

Despite the typical robustness and generality of the effect, however, the memorial advantage for information that has been generated versus only read does not always appear and/or can be eliminated under certain circumstances. For example, McNamara and Healy (1995a, 1995b, 2000) showed that a generation effect does not occur for the recall of answers to arithmetic problems unless retrieval strategies that reinstate procedures used at the time of study are induced at the time of test. And deWinstanley, E. L. Bjork, and R. A. Bjork (1996) demonstrated either the occurrence or nonoccurrence of a generation effect for the same studied materials by manipulating the match between the information strengthened during the generation task and the type of information required for optimal

performance on a subsequent memory test. These findings, respectively, have been interpreted as supporting the procedural account and the transfer-appropriate multifactor account of generation effects. These two accounts, though different in their details, are similar in their focus on the importance of the relationship between encoding and retrieval processes in the production of a generation advantage.

In the procedural account (Crutcher & Healy, 1989; McNamara & Healy, 1995a, 1995b, 2000), it is assumed that participants generating information at study, as opposed to reading it, are more likely to engage in procedures during study that can be reinstated during the retention test. Then, if these procedures are induced at the time of test, a generation advantage should occur and, conversely, if they are not, a generation advantage should not occur. In the transfer-appropriate multifactor account (deWinstanley et al., 1996)—an explanation that expanded on the two-factor (Hirshman & R. A. Bjork, 1988) and multifactor (McDaniel, Waddill, & Einstein, 1988) accounts of generation effects—it is assumed that generation effects occur when the nature of the generation task leads participants to focus their processing on the type of information to which a later retention test is sensitive and, conversely, when there is not a good match between these types of information, a generation effect should not occur. Thus, in these two accounts, both the occurrence of generation effects, and the lack thereof, depend upon the relationship between encoding and retrieval processes.

The occurrence of generation effects can also be influenced by instructions concerning how to encode to-be-read information. Begg, Vinski, Frankovich, and Holgate (1991), for example, found that a generation advantage

Correspondence concerning this article should be addressed to P. A. deWinstanley, Department of Psychology, 120 W. Lorain St., Oberlin, OH 44074-1086 (e-mail: patty.dewinstanley@oberlin.edu).

could be eliminated by instructing participants to use an imagery strategy rather than a pronunciation strategy to encode read items. Similarly, deWinstanley and E. L. Bjork (1997) found that a generation advantage could be eliminated by instructing participants concerning what type of memory test to expect and how to process information optimally in anticipation of that type of test. With the same set of study materials, participants given such instructions did not produce a generation advantage, whereas control participants did.

In the present study, we investigated whether the generation advantage could be eliminated without *explicitly* instructing participants how to process to-be-read items more effectively. Specifically, we wanted to know whether merely providing participants with the opportunity to experience or observe a generation advantage in their own performance on a memory test for items they had generated versus only read during a preceding study episode would be sufficient to induce them to process future to-be-read information in a more effective manner, a development that could be reflected in the elimination, or attenuation, of a generation advantage on subsequent memory tests. In other words, under certain circumstances, could participants discover for themselves how to engage in more effective processing during the acquisition of new information via reading?

In considering this possibility, it is important to note that past studies examining multiple attempts at generating and reading have found a small *increase* in the generation effect across tests (Gardiner & Arthurs, 1982; McNamara & Healy, 1995a), suggesting that—contrary to our speculation—the generation advantage would not be eliminated, or even attenuated, across tests; rather, it should increase. It is also important to note, however, that these previous studies critically differ from the type of situation with which the present hypothesis is concerned: namely, how the read-versus-generate encoding variable was manipulated. In the Gardiner and Arthurs (1982) study, the requirement to read versus generate was manipulated in a between-lists manner, and in the McNamara and Healy (1995a) study, the requirement to read versus generate was manipulated in a between-participants manner. Consequently, participants in these previous studies did not have the opportunity to experience the memorial consequence that resulted from generating versus reading within the context of the same memory test prior to future presentations of material to be encoded by reading or generating. In contrast, we ensured that participants in our first experiment would have such an experience prior to the presentation and testing of additional material. If, as speculated, such an experience would lead to better processing of future to-be-read information, a generation advantage should be attenuated or eliminated in subsequent tests.

EXPERIMENT 1A

Our primary goal in Experiment 1A was to determine whether the opportunity to experience for themselves—

within the context of the same test—the memorial advantage conferred upon items that participants had generated versus items they had only read during a previous study episode would be sufficient to lead them to adopt a more effective way of processing future to-be-read material. Our experimental strategy for testing this possibility was to present two paragraphs of text material to study, each containing critical to-be-generated and to-be-read items, with the presentation of each paragraph followed by its own memory test. Thus—before presentation of the second paragraph—participants would have engaged in both generating and reading critical items, and they would have had the opportunity to experience a generation advantage in their own performance on the memory test of those items. According to the present hypothesis, such an experience should be sufficient to induce more effective processing of future to-be-read items. If so, the to-be-read items in the second paragraph should benefit from this change, with the consequent attenuation, or even elimination, of a generation effect in the test of the second paragraph.

Method

Participants and Design. To fulfill a course requirement, 44 students of introductory psychology at Oberlin College participated in the present study. Of these students, 25 were assigned to study paragraphs on hypnosis and emotion, and 19 were assigned to study paragraphs on somatoform disorder and attitude formation.¹

Paragraphs were broken into phrases that were presented one at a time, with each containing either a critical to-be-generated word (generate-phrases) or a critical to-be-read word (read-phrases). These phrases were alternated across the presentation of the paragraph, resulting in a within-subjects manipulation of the requirement to generate versus read. A first and a second paragraph were presented to all participants in this manner, with the presentation of each followed by a memory test for the critical items presented in that paragraph, resulting in a within-subjects manipulation of paragraph presentation order. The crossing of these two variables thus created a 2 (encoding: generate vs. read) × 2 (paragraph order: first vs. second) within-subjects design.

Materials and Apparatus. The study materials were paragraphs taken from an introductory psychology textbook, and a given participant was either presented with one paragraph on attitude formation and another on somatoform disorder or with one paragraph on emotion and another on hypnosis. In addition, for both sets of paragraphs, the order in which the two paragraphs were presented and tested was counterbalanced across participants.

Each paragraph was separated into 16 phrases containing from 5 to 10 words, and the phrases were presented one at a time on the screen of a Gateway PC, with the first two phrases serving as primacy buffers. The 16 phrases alternated between generate-phrases and read-phrases, with half the paragraphs beginning with a generate-phrase and the other half beginning with a read-phrase. All words within a phrase appeared in black print except for the one critical item, which itself appeared in red print. To illustrate, consider the phrase, “The emotional or affective part,” and the word *affective* as the critical item. When “affective” appeared as a to-be-generated item, the participant would see, “The emotional or aff-ct-v- part,” with “aff-ct-v-” printed in red, and all other words printed in black; whereas, when “affective” appeared as a to-be-read item, the participant would see, “The emotional or affective part,” with “affective” in red letters, and all other words printed in black. In addition, across the participants presented with this phrase, half saw it as a generate-phrase and half saw it as a read-phrase, ensuring that

the same words appeared equally often as both types of critical items. During presentation of a paragraph, the participants wrote the solution to the word fragment, or they wrote the critical to-be-read word, on a separate page of a response booklet, turning the page after writing each word.

Presentation of each paragraph was followed by a brief distractor task and then a fill-in-the-blank test for the critical items, which contained all 16 phrases previously studied. The phrases were numbered, appeared on the same page, and in the same order as presented during study. In each phrase, however, one word was deleted and replaced by a blank line on which the participants were to write the missing word, and, across the 16 phrases, the two types of critical items were tested equally often in this manner. Furthermore, owing to counterbalancing procedures, the specific words that appeared and were tested as each type of item were the same across participants.

The fill-in-the-blank tests were contained in a single test booklet, and a message in bold letters told the participants not to turn each page until instructed to do so. The first page contained instructions about the experimental task. The second contained instructions for the distractor task that intervened between the presentation of each paragraph and its test, and this page also presented five words to be rehearsed as part of the distractor task. This second page was dark blue to prevent the participants from seeing through it to the third page, which contained the fill-in-the-blank test for the first paragraph with instructions for the test printed at the top. The next page, on which the participants could write the five words from the distractor task, was blank except for instructions printed at the top. Following these four pages, the pages relevant to the second paragraph began, and these were ordered in the same way: First, the dark blue page presenting the distractor task for the second paragraph; next, the page with the fill-in-the-blank test for the second paragraph; and, finally, the blank page for writing the words from the second distractor task.

Procedure. The participants were tested in small groups, with each person sitting individually at a computer terminal. After the participants were given the appropriate test and response booklets, the experimenter explained that they were in a study designed to test their memory and that they would be presented with two paragraphs, one at a time, with each paragraph followed by a test. They then read the instructions printed on the first page of the test booklet, which stated that: (1) they would be studying paragraphs presented one phrase at a time on the computer screen; (2) one item in each phrase would appear in red letters; (3) the red item could be a word fragment or an intact word spelled out in red letters; and (4) they should write the solution to each red fragment or the intact red word on a separate page in their response booklet, turning the page after writing in preparation for the next phrase. Examples of generate and read items were provided.

After their questions were answered, the participants were told to press any key to begin presentation of the phrases, each of which lasted for 10 sec. Following presentation of the last phrase, the experimenter collected the first response booklet while instructing the participants to turn to the first blue page in their test booklet and begin working on the task presented there, which was the distractor task. After engaging in this task for 2 min, the participants were told to turn the page and begin the fill-in-the-blank test. After 2 min, they were instructed to turn to the next page and were given 30 sec to write down the distractor items from the previous page. The participants then put aside their test booklet, located the second response booklet, and prepared to receive the second paragraph. After presentation of the 16 phrases of the second paragraph, they engaged in the second 2-min distractor task, then the 2-min fill-in-the-blank test for that paragraph, and, finally, the 30-sec distractor test. The participants were then debriefed and thanked for their participation.

Results and Discussion²

Figure 1 presents the mean percentages of correct answers obtained on the fill-in-the-blank test for the two

critical-item types as a function of paragraph presentation order. An examination of this figure indicates that a memory advantage for generated items occurred on the test of the first paragraph ($M = 63.27$ vs. $M = 41.95$ for generate and read items, respectively) but not on the test of the second paragraph ($M = 61.91$ vs. $M = 63.45$ for generate and read items, respectively), and this apparent pattern was confirmed by the finding of a significant interaction between encoding task (read vs. generate) and paragraph presentation order (first vs. second) [$F(1,43) = 11.54$, $MS_e = 498.23$, $p < .001$]. In addition, planned-comparison t tests revealed a significant increase in the memory performance of participants on read items from the test of the first paragraph to that of the second paragraph, whereas their performance on generate items did not differ significantly across the two tests [$t(43) = -4.83$, $p < .001$, and $t(43) = 0.24$, $p > .80$, respectively].

Thus, consistent with the present hypothesis, the generation advantage observed on the test of the first paragraph was eliminated on the test of the second paragraph because the recall of to-be-read items improved to the level of that for to-be-generated items. We believe that this significant increase in the recall of to-be-read items across paragraphs reflects the development of an improved processing strategy for such items as a consequence of the participants having had the opportunity to experience the memorial advantage for material they had generated rather than simply read in the test of the first paragraph. Having had that experience, the participants

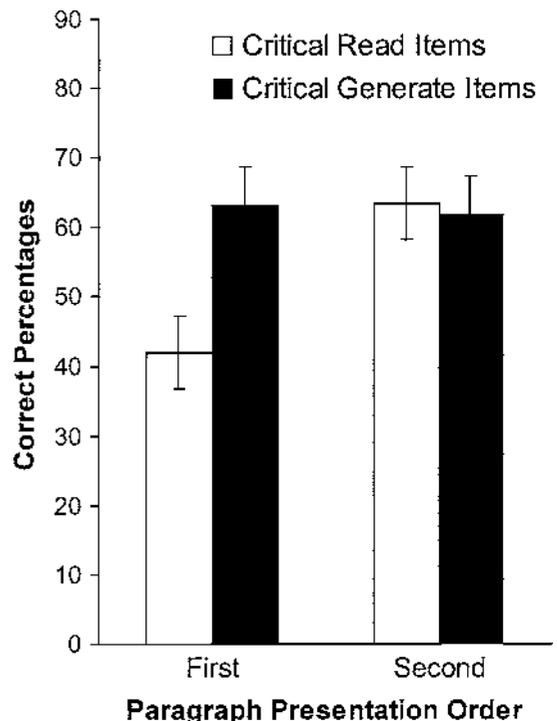


Figure 1. Percentages of correct answers for the two types of critical items as a function of paragraph order in Experiment 1A. Bars represent standard errors of the means.

improved their processing of future to-be-read items to the extent that they were able to recall such items as well as those they were required to generate, thereby eliminating a generation advantage on the second test.

EXPERIMENT 1B

In Experiment 1B, we tested the reliability and generality of the results of Experiment 1A by replicating its general experimental strategy while employing different materials and a different participant pool. In addition, in Experiment 1B, we included a task to ascertain whether the participants became consciously aware of the memorial advantage for generated items during the test for the first paragraph. Specifically, after they completed this test, we asked participants to respond to the question, "What did you notice about your performance on the previous memory test?" If, as previously speculated, the opportunity to experience the generation advantage in their own memory performance leads the participants to develop a more effective way of processing future to-be-read material, we thought that at least some of the participants' responses to this question would reflect such an awareness.

Method

Participants and Design. A total of 31 students from an educational psychology course at Oberlin College participated in the present study as volunteers. The participants were informed that they would be engaging in a classroom demonstration from which the results might be published, but in every other way it would be like any other classroom demonstration in which they may have participated in the past. Students were instructed that if they did not wish to participate, they could simply write "no response" in their booklets. Asking everyone to use a code name or symbol on their booklets and recall sheets preserved anonymity, thereby allowing students to refuse to participate without fear of penalty from the course instructor. No one in the class refused to participate. The design of Experiment 1B was in all other aspects the same as in Experiment 1A.

Materials. The study materials were paragraphs from an educational psychology textbook; one paragraph was on motivation and goal orientation, the other on Bloom's taxonomy of instructional objectives. Each paragraph was separated into 12 phrases containing from 7 to 16 words each. The phrases were presented one at a time on a page in a test booklet, with the first two phrases considered to be primacy buffers. All words within a phrase appeared in black print, except for the one critical item that was either to be generated or read, which itself appeared in bold black print and was underlined. A blank line was provided at the bottom of each page so that during the presentation of the paragraph, the participants could write down the solution to the word fragment in the case of a generate-phrase, or the critical to-be-read word in the case of a read-phrase, and turn the page only when instructed to do so by the experimenter.

The construction of the generate items and the counterbalancing of items across conditions was the same as in Experiment 1A. The tests were also constructed in the same way as in Experiment 1A. Unlike Experiment 1A, however, test booklets were not employed; rather, each test was distributed separately to the participants after they had completed the study phase. In addition, a single page with the question, "What did you notice about your performance on the previous memory test?" was stapled to the first fill-in-the-blank memory test.

Procedure. The experimenter told the participants, who were tested in one group, that they were in an experiment designed to test their memory. They were told that two paragraphs, broken into phrases, would be presented, with each paragraph followed by a test; that one item in each phrase would be printed in bold letters and underlined; and that they were to write the solution to each bolded fragment or the intact bolded word on the line provided at the bottom of their booklet, not turning the page until the experimenter instructed them to do so. An example of each type of item was written on the board at the front of the classroom.

The experimenter timed the presentation of each phrase with a stopwatch, instructing the participants to turn the page at 10-sec intervals. Following presentation of the last phrase, the experimenter collected the first booklet and distributed the test. The participants were given 2 min to fill in the 12 blanks. They were then instructed to turn to the attached page with the question, "What did you notice about your performance on the previous memory test?" and were given 2 min to respond to this question. The experimenter then collected the first test and distributed the second study booklet. After presentation of the 12 phrases of the second paragraph, the second study booklet was collected and the fill-in-the-blank test for that paragraph was distributed. Again, the participants were given 2 min to complete the 12 blanks on the test, after which it was collected. The participants were then thanked for their participation and debriefed in the following class meeting.

Results and Discussion

Despite the many procedural changes, as well as changes in the materials and the type of participants, Experiment 1B replicated the results of Experiment 1A. Figure 2 presents the mean percentages of correct answers obtained on the fill-in-the-blank test for the two types of critical items as a function of paragraph presentation order. As with Experiment 1A, a memory advantage for generate items occurred on the test of the first paragraph ($M = 49.68$ vs. $M = 33.55$ for generate and read items, respectively) but not on the test for the second paragraph ($M = 52.90$ vs. $M = 53.55$ for generate and read items, respectively). The interaction between encoding task (read vs. generate) and paragraph presentation order (first vs. second) was significant [$F(1,30) = 5.03$, $MS_e = 433.98$, $p < .05$]. Furthermore, planned-comparison t tests revealed the same pattern as that seen in Experiment 1A: a significant increase in recall for read items from the test of the first paragraph to that of the second paragraph, whereas recall for generate items did not differ significantly across the two tests [$t(30) = -3.70$, $p < .001$, and $t(30) = -0.55$, $p > .50$, respectively].

In addition to replicating the results of Experiment 1A, the present experiment allowed us to ascertain to what extent the participants became aware of a generation advantage in their performance on the first test. Responses to the question, "What did you notice about your performance on the previous memory test?" were coded into two categories on the basis of whether they reflected noticing a generation advantage. To illustrate, two examples of statements coded as "noticing a generation advantage" were: "I noticed that I mostly remembered the words I had to complete," and "I remembered the words I had to figure out"; whereas two examples of statements coded as "not noticing a generation advantage" were:

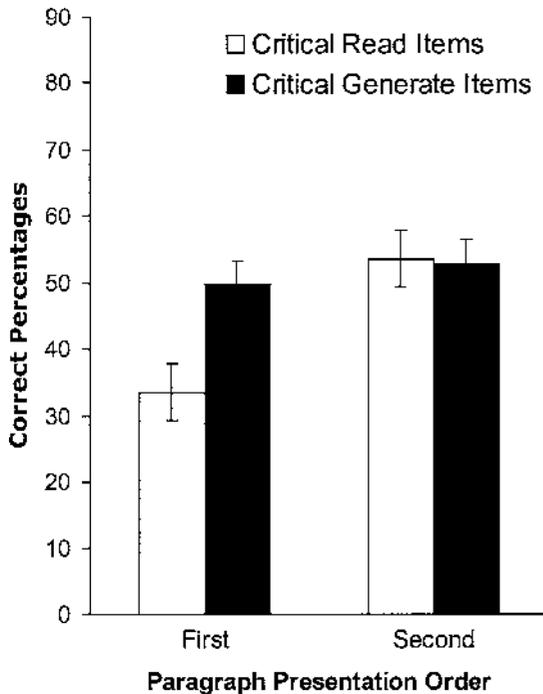


Figure 2. Percentages of correct answers for the two types of critical items as a function of paragraph order in Experiment 1B. Bars represent standard errors of the means.

“The later the underlined words appeared in the sentence the better I remembered it,” and “I improved in the second half of the questions.” In the statements coded as “noticing a generation advantage,” the participants referred to noticing better performance on words described in the following ways: completed items, items they had to figure out or fill in, words missing letters, words that were thought about as compared with just written, words that were generated, or words that were deciphered.

On the basis of this coding system, 17 of the 31 participants noticed a generation advantage after the first test and, for these 17 participants, 13 had shown a generation advantage, 3 had shown a tie between generate and read items, and 1 had shown a read advantage in performance on the first test. Of the remaining 14 participants, whose statements did not reflect having noticed a generation advantage in their performance on the first test, 6 had shown a generation advantage, 2 had shown a tie between generate and read items, and 6 had shown a read advantage in their performance on the first test. Of the 17 participants who had noticed the generation advantage, 13 either tied or showed a read advantage in their performance on the second test, which, according to a chi-square test is more than would be expected by chance [$\chi^2(1) = 4.77, p < .05$]. Of the 14 participants who did not comment on a generation advantage, 8 either tied or showed a read advantage, and 6 showed a generation advantage in their performance on the second test, which is no more than would be expected by chance [$\chi^2(1) = 0.29, p < .50$]. Thus, on

the basis of their written comments, it would seem that just over half of the participants in the present experiment became aware of the memorial advantage of generating as compared with reading after the first test and, of these, most then changed the way they processed to-be-read items presented in the second paragraph, as reflected by their improved performance for such items on the second test.

EXPERIMENT 2

The lack of a multitrial generation effect observed in the present Experiments 1A and 1B stands in marked contrast to the results of Gardiner and Arthurs (1982), who did observe a generation effect across multiple trials. As previously mentioned, however, in their research the generate-versus-read encoding variable was manipulated within participants but between lists; whereas, in the present research, this encoding variable was manipulated both within participants and within paragraphs. Thus, the participants in Gardiner and Arthurs’s study—unlike the present ones—would not have had the opportunity to experience the memorial consequences resulting from generating versus reading within the context of the same memory test prior to future presentations of material to be encoded by generating or reading. If, as we believe, this experience led the participants in the present research to develop a more effective strategy for processing future to-be-read information with the consequential elimination of a multitrial generation effect, it should be possible to maintain such an effect by denying the participants this experience.

To test this possibility, we manipulated the requirement to generate versus read between paragraphs—rather than within paragraphs—in Experiment 2. By making this change, the participants would not have the opportunity to experience the memorial consequences of generating versus reading on the same memory test prior to being presented with additional material to read or generate. Thus—in contrast to the results of Experiments 1A and 1B, but consistent with our explanation of those results—a generation advantage should be maintained across paragraphs.

Method

Participants and Design. To fulfill a course requirement, 68 students of introductory psychology at Oberlin College participated in the present experiment. The participants were presented with the somatoform disorder and attitude formation paragraphs from Experiment 1A, broken into phrases, with the critical to-be-generated or to-be-read items highlighted in red print. Unlike Experiment 1A, however, the critical items appearing in the first paragraph presented to any one participant were only to-be-generated items or to-be-read items. Then, for the second paragraph, the encoding task was switched for each participant. Thus, although all participants engaged in both types of encoding, each participant switched from the generate-encoding to the read-encoding task or from the read-encoding to the generate-encoding task from the first to the second paragraph, with each paragraph followed by its own fill-in-the-blank test. The combination of these conditions thus produced a 2×2 mixed-subjects design, with paragraph order (first vs. second)

manipulated within participants and encoding-task order (read-to-generate vs. generate-to-read) manipulated between participants. Of the 68 participants, 37 were randomly assigned to the read-to-generate order and 31 to the generate-to-read order.

Materials and Apparatus. Materials and equipment were the same as in Experiment 1A, with the exception that only the somatoform-disorder and attitude-formation paragraphs were used, and the phrases within a given paragraph were either all generate-phrases or all read-phrases. As in Experiment 1A, the critical item appeared as a word fragment printed in red in a generate-phrase and as an intact word printed in red in a read-phrase, while all other words were printed in black.

Procedure. The procedure was the same as in Experiment 1A except that, for the first paragraph, the participants were only given instructions and examples appropriate for the encoding task to be used for that paragraph; then, after the test for the first paragraph, they were given instructions and examples for the encoding task to which they were switching for the second paragraph.

Results and Discussion

Figure 3 presents the mean percentages of correct answers obtained on the fill-in-the-blank test for the critical items presented in the first versus the second paragraph. In examining the pattern of performance for these items, it is important to keep in mind that the critical generate-to-read items were generated items in the first paragraph and read items in the second paragraph. Conversely, the critical read-to-generate items were read items in the first paragraph and generated items in the second paragraph.

An examination of Figure 3 indicates that an advantage for generated items occurred in the tests of both the

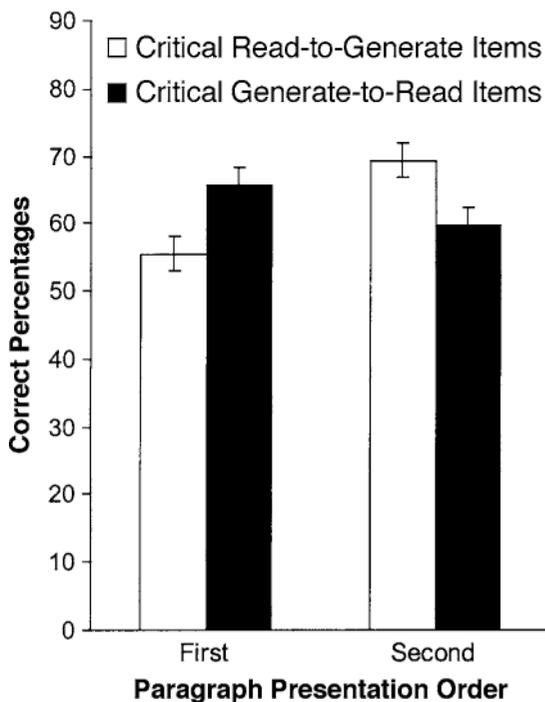


Figure 3. Percentages of correct answers for the two types of critical items as a function of encoding task and paragraph order in Experiment 2. Bars represent standard errors of the means.

first and second paragraphs. And indeed, this apparent result was confirmed by the outcome of a mixed analysis of variance (ANOVA), which revealed both a generation advantage [$F(1,66) = 6.38, MS_e = 520.03, p < .01$], and that the advantage for generated items occurred in the absence of an order effect [$F(1, 66) = 1.01, MS_e = 520.03, p > .30$]. Moreover, planned comparisons revealed memory performance to be significantly better for generated than for read items both in the test of the first paragraph ($M = 65.58$ and $M = 55.46$, respectively) and in the test of the second paragraph ($M = 69.32$ and $M = 59.61$, respectively) [$t(66) = 1.64, p < .05$; and $t(66) = 1.66, p < .05$, respectively]. Also, average performance on the tests for the first and second paragraphs did not differ significantly [$t(67) = -1.19, p > .10$].

Thus, as predicted, when the participants were denied the opportunity to observe the memorial consequences of generating versus reading within the same test before the presentation of additional information to learn, a generation advantage was obtained for both the first and second paragraphs, lending strong support to the notion that this experience led the participants of Experiments 1A and 1B to adopt an improved processing strategy for read items between the first and second paragraphs. Denied this experience, the participants of Experiment 2, even those who did engage in generation during the first paragraph, did not develop a better processing strategy for the to-be-read words when switched from generating to reading in the second paragraph. Hence, the typical memory advantage for generate versus read items was also observed in the test for the second paragraph; that is, as predicted, a multitrial generation effect was obtained in Experiment 2.

It should also be noted that the results of Experiment 2 rule out an explanation for the results of Experiments 1A and 1B in terms of the participants having simply learned to pay special attention to words appearing in red or bold print—that is, their having learned that those words would later be tested and so the participants ought to pay special attention to them in the second paragraph. If such an attention-based account correctly explained the elimination of a generation effect across paragraphs in Experiments 1A and 1B, the generation effect should also have been eliminated across paragraphs in Experiment 2, as these latter participants had the opportunity to make this same observation about the items printed in red.

EXPERIMENT 3

As conjectured, when the participants were denied the opportunity to observe the memorial consequences of generating versus reading prior to the presentation of additional material—owing to the between-paragraphs manipulation of the read/generate variable in Experiment 2—a generation advantage was maintained across paragraphs, consistent with our hypothesis that the elimination of the generation advantage in Experiments 1A and 1B occurred because the participants changed to a more effective

tive strategy for processing to-be-read items after experiencing the memory benefits of generating relative to reading. Although we see this experience as playing a critical role in leading participants to make this change, it is possible that it does not play such a role. Perhaps, instead, participants just felt dissatisfied with their ability to recall read items on the test of the first paragraph—a feeling that arose independently of any experience with generating, but which nonetheless led them to adopt a more effective processing strategy for the subsequently presented read items. The results of Experiment 2 cannot rule out such a possibility because—even if a sense of dissatisfaction with their ability to recall the critical read items presented in the first paragraph would have induced the participants to adopt a more effective processing strategy for such items in a second paragraph—the switch from a read to a generate encoding task between paragraphs made it impossible for them to demonstrate any such change.

Experiment 3 tested this possibility by employing the same multiple study–test procedure of Experiments 1A, 1B, and 2, but manipulating the requirement to generate versus read between participants rather than between paragraphs. In this situation, we would expect to see a typical generation effect on the test for the first paragraph—that is, better memory performance on critical items by the participants who generated them than by the participants who read them during study of the first paragraph. If, however, something like a general feeling of dissatisfaction with one's performance on read items presented in a first study episode is sufficient to cause a change to a more effective processing strategy for future to-be-read items, we should see a reduction or elimination of a generation advantage on the test for the second paragraph, as was observed in Experiments 1A and 1B. If, rather, as we believe, the opportunity to experience the memory benefits of generation relative to reading is critical for inducing such a change, the participants in Experiment 3 should not change their processing strategy for the to-be-read items between the first and second paragraphs, and a generation advantage should be seen on both tests. In addition to supporting the present explanation of why a generation advantage was not observed on the second test in Experiments 1A and 1B, the observation of a multi-trial generation effect in Experiment 3 would be consistent with the research of McNamara and Healy (1995a), who manipulated the requirement to read versus generate in a between-subjects design and found a generation advantage across multiple memory tests.

Method

Participants and Design. To fulfill a course requirement, 93 students of introductory psychology at Oberlin College participated in the present study. Of these, 46 were randomly assigned to the generate condition, and 47 were randomly assigned to the read condition. In both encoding conditions, the participants were presented with the same two paragraphs used in Experiment 2, with the presentation of each followed by a memory test for the critical items presented in that paragraph. Unlike Experiment 2, however, both of

the paragraphs presented to any one participant either contained only read-phrases or only generate-phrases; that is, the participants were not switched from one type of encoding task to the other across paragraphs. Thus, the design of Experiment 3 was a 2×2 mixed-subjects design, with type of encoding (generate vs. read) manipulated between participants and paragraph presentation order (first vs. second) manipulated within participants.

Materials and Apparatus. All materials and equipment used were the same as those employed in Experiment 2, except that the phrases within the two paragraphs presented to a given participant were either all generate-phrases or all read-phrases.

Procedure. The procedure was the same as those used for Experiments 1A and 2, except that the participants were only given instructions appropriate for the encoding-task condition to which they had been assigned. Then, after administration of the first test, they were simply reminded of this same encoding task and given a second example before presentation of the second paragraph.

Results and Discussion

Figure 4 presents the mean percentages of correct answers obtained on the fill-in-the-blank test for the critical to-be-generated and to-be-read items as a function of paragraph presentation order. These results show that a memory advantage for generated items over read items occurred not only in the tests for the first paragraph ($M = 62.00$ vs. $M = 47.94$ for generate and read items, respectively) but also in the tests for the second paragraph ($M = 70.78$ vs. $M = 59.64$ for generate and read items, respectively). And, indeed, a two-way mixed-subjects ANOVA revealed a significant overall generation ad-

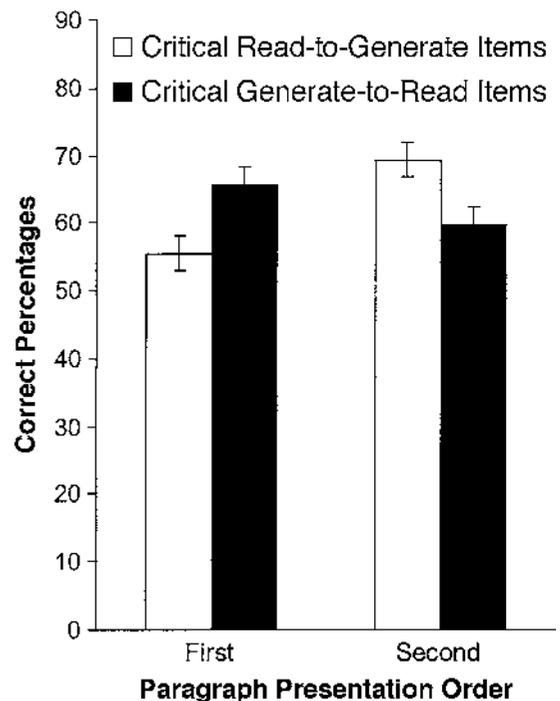


Figure 4. Percentages of correct answers for the two types of critical items as a function of encoding task and paragraph presentation order in Experiment 3. Bars represent standard errors of the means.

vantage and no significant interaction between encoding condition and paragraph presentation order [$F(1,91) = 6.78$, $MS_e = 1,051.11$, $p < .01$; $F(1,91) = 2.61$, $MS_e = 257.96$, $p > .10$, respectively]. Furthermore, planned comparison t tests confirmed that a significant generation advantage was obtained on the tests for both the first and the second paragraphs [$t(91) = 2.53$, $p < .05$; $t(91) = 2.37$, $p < .05$, respectively].

An examination of Figure 4 also indicates that the participants' memory for critical items improved across the two paragraphs, and this apparent improvement was confirmed by the finding of a significant overall increase in memory performance from the first to the second paragraph [$F(1,91) = 15.12$, $MS_e = 831.44$, $p < .001$], but the increase in memory performance across paragraphs did not differ between the two encoding conditions [$F(1,91) = 0.17$, $MS_e = 831.44$, $p > .60$]. Hence, although the participants in the read-encoding condition did show an improvement in their memory for critical to-be-read items from the first to the second paragraph, their degree of improvement is not different from that shown by the participants in the generate-encoding condition for critical to-be-generated items, nor was it sufficient to eliminate the memorial advantage of the generated items in the tests of the second paragraph. Thus, in contrast to the results of Experiments 1A and 1B, the improved memory performance on critical to-be-read items across paragraphs in Experiment 3 would appear to reflect the influence of a more general learning strategy that affected the performance of the participants in both encoding conditions, rather than reflecting a switch to a more effective processing strategy specifically for read items from the first to the second paragraph.

GENERAL DISCUSSION AND CONCLUSIONS

Stated most generally, the present research has explored whether participants could discover for themselves how to process or encode to-be-read information more effectively, or, put slightly differently, discover for themselves how to become more effective readers. Specifically, we hypothesized that allowing participants to experience in their own performance the memorial benefit conferred upon items they had generated versus items they had only read during a previous learning episode would induce them to develop a more effective strategy for processing subsequently presented to-be-read information, which would be reflected in the improved recall of such information and the possible elimination of a generation advantage on future memory tests. To investigate this possibility, we gave the participants in Experiments 1A and 1B such an experience between study of a first and second paragraph and, consistent with the hypothesis, although a generation advantage was observed for the first paragraph, recall of to-be-read items presented in the second paragraph then improved to the level of that for generated items, eliminating a generation advantage.

Furthermore, the results of Experiments 2 and 3 strongly implicated that such an experience was critical to the development of an improved processing strategy for future to-be-read items. When participants were denied this experience, either because the read/generate encoding variable was manipulated between paragraphs or between participants, the recall advantage for generated items was maintained across tests.

Although other studies have shown that generation effects can be eliminated when participants improve their processing strategy for to-be-read items (e.g., Begg et al., 1991; deWinstanley & E. L. Bjork, 1997), an important difference exists between these previous findings and the present ones. Namely, in the previous studies, generation effects were only eliminated when participants were explicitly instructed in how to improve their processing of to-be-read items; whereas, in the present research, generation effects were eliminated simply by giving participants the opportunity to observe the consequent memory advantage for generated versus read items in their own recall performance. Thus, as the present findings demonstrate, not only can generating produce a memorial advantage over reading when learning new information; it can, under certain circumstances, also lead to the spontaneous development of processing strategies for future to-be-read material that render it just as well learned as generated material.

In drawing these conclusions from the present research, we do not mean to imply that participants could not be induced to improve their processing of to-be-read information in other ways in the present learning situation. As in the studies of Begg et al. (1991) and deWinstanley and E. L. Bjork (1997), it might well be possible to induce improved processing of to-be-read information in the present situation with explicit instructions. In addition, we do not mean to imply that the present experience is unique in its ability to induce the spontaneous development of more effective strategies for the processing of information encoded by reading. In a study by McNamara and Healy (1995a), for example, in which read and generate participants were given multiple trials to learn a list of pairs, many read participants spontaneously developed improved processing strategies, and then performed as well as generate participants on a later retention test. McNamara and Healy's (1995a) study thus provides an example of another situation—namely, the acquisition of information across multiple trials—in which learners are sometimes able, on their own, to develop more effective ways of processing to-be-read information. It is thus interesting to speculate whether the read-only participants in the present situation, who did not develop an improved processing strategy for to-be-read words, might eventually have done so given multiple paragraphs and tests. Possibly, however, multiple training trials involving tests on the same to-be-learned material—as opposed to multiple trials involving the learning and testing of new material—is critical for read-only participants to develop improved processing strategies.

Another subject for future exploration is the question of just what participants in the present learning situation did to improve their processing of to-be-read items. One possibility might be as follows. During study of the first paragraph, participants used contextual information provided by other words—as well as the presented letters—to help complete the to-be-generated items and then used such information again to help recall those same items during the test. The use of such a strategy, first to generate and then to recall the items generated, could underlie the generation advantage for the first paragraph—an explanation that would be consistent with both the procedural account (Crutcher & Healy, 1989; McNamara & Healy, 1995a, 1995b, 2000) and the transfer-appropriate version (deWinstanley et al., 1996) of the multifactor accounts of generation effects (Hirshman & R. A. Bjork, 1988; McDaniel et al., 1988). That is, it would be the match between the information that participants strengthened during generation and the information needed to perform well on a later test, or the ability to reinstate during test the cognitive procedures used during study, that led to the initial generation advantage. Then, having noticed both their superior recall of generated items and their use of contextual information in recalling them, participants were led to attend to such contextual information during study of the second paragraph, whether the critical item was to be read or generated.

Whether improving participants' strategy for processing future to-be-read items in this or some other way, however, the results of the present Experiments 1A, 1B, and 2 can be viewed as analogous to those previously obtained by McNamara and Healy (1995b). Whereas these investigators found no generation advantage for the recall of answers to addition problems in contrast to multiplication problems when presented separately, they found a generation advantage for both when presented together. To explain this difference, McNamara and Healy (1995b) argued that only when participants received both problems did they realize that a retrieval strategy used to recall answers to generated multiplication problems could also be used to recall answers to generated addition problems. Similarly, in the present research, it was only when participants were asked to recall both generated and read items within the same test that they extended any improved processing strategies learned for generated items to read items.

Interestingly, the explanation we are proposing to account for the present results relates in certain respects to one recently proposed by Sahakyan and colleagues (e.g., Sahakyan, Delaney, & Kelley, 2004; Sahakyan & Kelley, 2002) to account for certain findings observed in connection with an entirely different phenomenon—that of directed forgetting. A frequent finding in studies of directed forgetting is the superior recall of a second-presented list (List 2) by participants told to forget a first-presented list (List 1) as compared with the recall of List 2 by participants told to remember List 1 after its presentation. (For a complete review of this phenomenon

and related findings, see R. A. Bjork, 1972, 1989; Johnson, 1994; and MacLeod, 1998.) One explanation proposed for the superior List-2 recall of forget participants is in terms of retrieval inhibition (e.g., E. L. Bjork & R. A. Bjork, 1996; R. A. Bjork, 1989) according to which the instruction to forget List 1, followed by the presentation of List 2 to learn in its stead, initiates a process that inhibits subsequent retrieval access to List 1. With its retrieval access inhibited, the proactive interference of List 1 on the recall of List 2 is decreased or eliminated, leading to its better recall by forget versus remember participants. Alternatively, Sahakyan and colleagues (Sahakyan et al., 2004; Sahakyan & Kelley, 2002) have proposed an explanation for this finding in terms of differential encoding of List 2. Whereas both types of participants are assumed to encode List 1 in the same way, typically using an ineffective maintenance rehearsal strategy, it is assumed that forget participants, being told to forget List 1, then reflect upon their encoding of it, coming to realize the ineffectiveness of their encoding strategy and to switch to a better one for List 2. In contrast, remember participants are thought to be less likely to engage in such reflection and thus do not switch to a better encoding strategy for List 2. Consistent with this idea, Sahakyan et al. (2004) found the superior List-2 recall by forget versus remember participants to disappear when both were given an opportunity to recall List 1 before being told to forget or remember it, or when asked to evaluate their likely performance on List 1 before presentation of List 2.

Thus, the Sahakyan et al. (2004) explanation for why forget participants improve their recall of List 2 and our explanation for why Experiments 1A and 1B participants improve their recall of to-be-read items are similar in at least one respect; namely, they both propose an improved encoding strategy that underlies the observed recall improvements. At least one important difference between these two situations and proposed explanations should be highlighted, however. In the Sahakyan et al. explanation, it must be assumed that the forget participants are somehow able to realize the need for a better processing strategy for List 2 without the experience of first recalling List 1. But in the present paradigm, participants are apparently unable to realize the relative ineffectiveness of reading versus generating without actually experiencing the differential performance consequences of these two types of encoding strategies.

In this respect, the present results and explanation are consistent with previous research indicating that individuals are typically unable to judge the efficacy of a given processing strategy during its execution and do not switch from a less- to a more-effective strategy without an opportunity to experience their relative effectiveness (e.g., Brigham & Pressley, 1988; Dunlosky & Hertzog, 2000; Shaughnessy, 1981). In particular, our results and explanation seem consistent with the framework of knowledge updating proposed by Dunlosky and Hertzog (2000) in which individuals are assumed not to update their

knowledge about the effectiveness of different encoding strategies unless they become aware of the consequential performance differences for information processed with different strategies. In line with that assumption, only participants in the present learning situation who experienced the relative performance consequences of the two encoding strategies then went on to adopt a more effective processing strategy for to-be-read items.

In addition to being consistent with the knowledge-updating framework of Dunlosky and Hertzog (2000), the present findings are consistent with the characterization of the learner emerging from research on how individuals monitor their level of comprehension during study (e.g., Benjamin & R. A. Bjork, 1996; Benjamin, R. A. Bjork, & Schwartz, 1997; R. A. Bjork, 1999; Dunlosky & Nelson, 1994; Koriat, 1997, 1998; Mazzoni & Nelson, 1995; Metcalfe & Shimamura, 1994; Nelson & Narens, 1990). Researchers studying such monitoring processes (e.g., R. A. Bjork, 1999; Jacoby, R. A. Bjork, & Kelley, 1994) have argued that the readings individuals take of their level of comprehension during a learning task can be as important as their actual comprehension, in part, because such readings can highly influence how individuals allocate their learning resources. Problematically, however, learners are not very accurate in taking such readings and often suffer from what has been called illusions of comprehension or competence (R. A. Bjork, 1999; Jacoby et al., 1994). Learners, for example, can be fooled into thinking that their state of knowledge or competence is greater than it is by conditions of the learning environment (such as massed practice) that enhance or sustain their performance during learning, but actually impair their long-term retention and/or transfer (R. A. Bjork, 1999; Christina & R. A. Bjork, 1991; Schmidt & R. A. Bjork, 1992; Simon & R. A. Bjork, 2001). Or, as Koriat (1997) has argued, learners can suffer from illusions of competence because they are relatively insensitive to the presence of factors in the learning environment that actually enhance later memory performance (such as repeated presentations) and overly sensitive to factors that do not necessarily lead to enhanced later memory performance (such as the perceived association between a cue and a target when both are present during the study trial).

The notion that learners can be insensitive to the effectiveness of certain learning enhancing factors can be applied to the present research to explain the results obtained in Experiment 2 versus Experiments 1A and 1B. Accordingly, the participants in Experiment 2, who were switched from a generating to a reading task between paragraphs, would be seen as being insensitive to the mnemonic benefit of generating over reading. Although tested after the first paragraph on the items they had generated, they were apparently unable to appreciate that their recall of these items had been improved over what it would have been had they only read them. Thus, they did not change how they processed the to-be-read items during study of the second paragraph. In contrast, the

participants in Experiments 1A and 1B, experiencing both the processing differences between generating and reading, and the resulting differences in recall, apparently enhanced their sensitivity to the mnemonic ineffectiveness of typical reading and, as a consequence, upgraded their processing of the subsequently presented to-be-read items. Some aspects of the present results are thus consistent with the assertion that learners are often *not* sensitive to procedures that contribute to better long-term retention, while other aspects have the more optimistic implication that—given the right types of experiences—learners can become more sensitive to such procedures and consequently improve their ability to learn new materials in the future.

In conclusion, as demonstrated by each of the present experiments, generation advantages extend to the learning of materials contained in textbooks. Thus, from an applied perspective, the picture painted by the present findings seems promising, although exactly how to incorporate generation strategies into educational materials to maximize learning remains a subject for future exploration. One suggestion from the present findings for such endeavors, however, is that students may need to experience for themselves the differential advantage of generation in the context of an actual test for them to realize its effectiveness as an encoding strategy. Learning exercises that just require generation—although likely to produce better memory performance for the generated items—are not likely to induce students to adopt more effective strategies for reading. Possibly, as suggested by the present research and that of Dunlosky and Hertzog (2000), allowing students to actually experience the performance consequences of differentially effective encoding processes may be necessary before they can appreciate their different degrees of efficacy and make profitable use of such knowledge in their future study activities.

REFERENCES

- BEGG, I., VINSKI, E., FRANKOVICH, L., & HOLGATE, B. (1991). *Generating makes words memorable, but so does effective reading*. *Memory & Cognition*, *19*, 487-497.
- BENJAMIN, A. S., & BJORK, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. Reder (Ed.), *Implicit memory and metacognition: The 27th Carnegie Symposium on Cognition* (pp. 309-338). Mahwah, NJ: Erlbaum.
- BENJAMIN, A. S., BJORK, R. A., & SCHWARTZ, B. L. (1997). The mis-measure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55-68.
- BJORK, E. L., & BJORK, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness & Cognition*, *5*, 176-196.
- BJORK, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, DC: Winston & Sons.
- BJORK, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Erlbaum.
- BJORK, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII—Cognitive regulation of performance: Interaction of theory and application*. Cambridge, MA: MIT Press.

- BRIGHAM, M. C., & PRESSLEY, M. (1988). Cognitive monitoring and strategy choice in younger and older adults. *Psychology & Aging*, *3*, 249-257.
- CHRISTINA, J. F., & BJORK, R. A. (1991). Optimizing long-term retention and transfer. In D. Druckman & R. A. Bjork (Eds.), *In the mind's eye: Enhancing human performance* (pp. 23-56). Washington, DC: National Academy Press.
- CRUTCHER, R. J., & HEALY, A. F. (1989). Cognitive operations and the generation effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 669-675.
- DEWINSTANLEY, P. A. (1995). A generation effect can be found during naturalistic learning. *Psychonomic Bulletin & Review*, *2*, 538-541.
- DEWINSTANLEY, P. A., & BJORK, E. L. (1997). Processing instructions and the generation effect: A test of the multifactor transfer-appropriate processing theory. *Memory*, *5*, 401-421.
- DEWINSTANLEY, P. A., BJORK, E. L., & BJORK, R. A. (1996). Generation effects and the lack thereof: The role of transfer-appropriate processing. *Memory*, *4*, 31-48.
- DUNLOSKY, J., & HERTZOG, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology & Aging*, *15*, 462-474.
- DUNLOSKY, J., & NELSON, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory & Language*, *33*, 545-565.
- GARDINER, J. M., & ARTHURS, F. S. (1982). Encoding context and the generation effect in multitrial free-recall learning. *Canadian Journal of Psychology*, *36*, 527-531.
- HIRSHMAN, E. L., & BJORK, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 484-494.
- JACOBY, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning & Verbal Behavior*, *17*, 649-667.
- JACOBY, L. L., BJORK, R. A., & KELLEY, C. M. (1994). Illusions of comprehension, competence, and remembering. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 57-81). Washington, DC: National Academy Press.
- JOHNSON, H. (1994). Processes of successful intentional forgetting. *Psychological Bulletin*, *116*, 274-292.
- KORIAT, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *127*, 349-370.
- KORIAT, A. (1998). Illusions of knowing: The link between knowledge and metaknowledge. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 16-34). London: Sage.
- LAWSON, M. J., & CHINNAPPAN, M. (1994). Generative activity during geometry problem solving: Comparison of the performance of high-achieving and low-achieving high school students. *Cognition & Instruction*, *12*, 61-93.
- MACLEOD, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Hillsdale, NJ: Erlbaum.
- MAZZONI, G., & NELSON, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 1263-1274.
- MCDANIEL, M. A., WADDILL, P. J., & EINSTEIN, P. J. (1988). A contextual account of the generation effect: A three-factor theory. *Journal of Memory & Language*, *27*, 521-536.
- MCNAMARA, D. S., & HEALY, A. F. (1995a). A generation advantage for multiplication skill training and nonword vocabulary acquisition. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity*. Thousand Oaks, CA: Sage.
- MCNAMARA, D. S., & HEALY, A. F. (1995b). A procedural explanation of the generation effect: The use of an operand retrieval strategy for multiplication and addition problems. *Journal of Memory & Language*, *34*, 399-416.
- MCNAMARA, D. S., & HEALY, A. F. (2000). A procedural explanation of the generation effect for simple and difficult multiplication problems and answers. *Journal of Memory & Language*, *43*, 652-679.
- METCALFE, J., & SHIMAMURA, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.
- NELSON, T. O., & NARENS, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125-141). San Diego: Academic Press.
- PESTA, B. J., SANDERS, R. E., & MURPHY, M. D. (1999). A beautiful day in the neighborhood: What factors determine the generation effect for simple multiplication problems? *Memory & Cognition*, *27*, 106-115.
- PEYNIRCIOĞLU, Z. F., & MUNGAN, E. (1993). Familiarity, relative distinctiveness, and the generation effect. *Memory & Cognition*, *21*, 367-374.
- SAHAKYAN, L., DELANEY, P. F., & KELLEY, C. M. (2004). Self-evaluation as a moderating factor of strategy change in directed forgetting benefits. *Psychonomic Bulletin & Review*, *11*, 131-136.
- SAHAKYAN, L., & KELLEY, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 1064-1072.
- SCHMIDT, R. A., & BJORK, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207-217.
- SHAUGHNESSY, J. J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. *Journal of Verbal Learning & Verbal Behavior*, *20*, 216-230.
- SIMON, D., & BJORK, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 907-912.
- SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 592-604.
- WITTROCK, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, *24*, 345-376.

NOTES

1. Experiment 1 was originally conducted in two parts, one with the somatoform disorder/attitude formation paragraphs and one with the hypnosis/emotion paragraphs, across two different semesters. The results, on the basis of these two sets of materials, did not differ significantly and have been combined in the present report.

2. The failure to generate items at study was extremely low (less than 1%); thus, all analyses for the present three experiments are uncorrected with respect to generation performance at study.