

# The effects of generation on item and order retention in immediate and delayed recall

DANIEL J. BURNS

*Union College, Schenectady, New York*

and

ELISABETH T. CURTI and JAMES C. LAVIN

*Lafayette College, Easton, Pennsylvania*

Recent research has shown that generating words from fragments, relative to simply reading them, inhibits processing of order information. Nairne, Riegler, and Serra (1991) showed that this reduction in processing of order information leads to deficits in recall performance. In three experiments, we generally replicate Nairne et al.'s results and demonstrate that the deficit in recall for the generated items is dependent on the easy distractor task and the relatively short (30-sec) retention interval they used. When a difficult distractor task was used, generating produced a deficit in amount of order information processed, but actually facilitated recall when recall was delayed 80 sec. The results are consistent with the hypothesis that generating words inhibits order processing, but they do not support the contention that the reduction in order processing is responsible for the deficit in recall that is sometimes observed for the generated items. The importance of the item-order distinction in explaining the generation effect, as well as the role of the item-order distinction in the long-term-memory arena, is questioned.

The memorial benefits of generating items from word fragments, relative to simply reading the items, have been well documented in within-list designs (e.g., Slamecka & Graf, 1978). More recent studies using between-lists designs have shown that generation is not always beneficial to long-term retention performance (e.g., Begg & Snider, 1987; Burns, 1990; Slamecka & Katsaiti, 1987). In fact, several studies have demonstrated that generating items, relative to simply reading them, can be detrimental to recall performance (Burns, 1992; Nairne, Riegler, & Serra, 1991; Schmidt & Cherry, 1989; Serra & Nairne, 1993).

Some of these recent studies point to a processing tradeoff explanation of the memorial consequences of generation. Generating induces the processing of item-specific and/or stimulus-response features (the shared features of the stimulus and response members of paired associates). However, the processing induced by generating is achieved at the expense of processing relational information (e.g., serial order information) (Burns, 1992; Nairne et al., 1991). For example, Burns (1992, Experiment 1C) showed that simply reading the items sometimes produced higher free recall and greater correspondence between input serial position and recall output order than generating the items. It was concluded that generation inhibited the processing of order information.

Nairne et al. (1991) used a more direct measure of relational (i.e., order) processing in their experiments. For example, in their Experiment 2, subjects were given 24 lists, each consisting of eight read items or eight generate items (word fragments). A 30-sec distractor task was given immediately after the presentation of each list, and was followed either by a free-recall test or an order-reconstruction test (subjects were presented all eight items and were asked to recreate the input order of the list). The simply read items produced both higher recall and higher reconstruction performance. It was concluded that generation impeded the processing of serial order information, which hindered long-term free-recall performance.

Serra and Nairne (1993) recently replicated Nairne et al.'s (1991) results in an incidental learning task. In addition, they showed that both the free-recall and the order-reconstruction advantage for the read items was eliminated when a mixed-list design was used rather than a between-lists design.

Burns (1986, Experiment 4) performed an experiment quite similar to those of Nairne et al. (1991). He presented subjects with 12 lists of six items, with each list being recalled either immediately, or after 40 or 80 sec of distractor activity. The items on half the lists were read items and on the other half they were generate items (word fragments). Burns found that the simply read items were recalled better than the generated items during immediate recall. However, the generated items were actually recalled better when an 80-sec retention interval was given. Thus, Burns's results are at odds with the results obtained by Nairne et al. Specifically, Burns found a recall advantage for the simply read items only in immediate recall,

---

The authors thank James Nairne, Robert Proctor, Norman Slamecka, and an anonymous reviewer for their comments on an earlier version of this manuscript. Correspondence concerning this article should be sent to D. J. Burns, Department of Psychology, Union College, Schenectady, NY 12308.

not when recall was delayed. Conversely, Nairne et al. found a recall advantage when recall was delayed.

An assessment of whether the additional order processing given to the read items improves long-term free recall is theoretically important to the general study of human memory, as well as to researchers interested in the generation effect. It is generally accepted that both item and order information contribute to free-recall performance in short-term retention tasks. However, much less is known about the role of item and order information in long-term retention tasks. Nairne and his colleagues (e.g., Nairne, 1990) have suggested that temporal order information may be an important factor in long-term retention tasks, and that the item-order distinction might prove to be a useful dichotomy in the long-term-memory arena. Similar arguments have been made in prose-processing research (Einstein, McDaniel, Bowers, & Stevens, 1984).

Nairne et al.'s (1991) study provides evidence that processing order information facilitates long-term recall, thus supporting the hypothesis that both item and order information improve long-term retention performance. Burns's (1986) Experiment 4 contradicts Nairne et al.'s finding. Thus, it is possible that order information does not facilitate long-term free recall. It seems important to understand the cause of the difference in results between the two studies, because it may help us to understand the role of order information in long-term memory.

There were several small procedural differences between Burns's (1986) experiment and the study by Nairne et al. (1991) that could have led to the difference in results. One difference that seemed especially important was the type of distractor task used during the retention interval. Nairne et al. presented subjects with a single digit every 1 sec and required them to determine whether the digit was odd or even. Burns, on the other hand, required subjects to study and recall lists of six two-digit items during the distractor period. It seemed to us that recalling lists of two-digit items was a much more difficult task than the odd-even digit task. In fact, Burns (1986) selected the number-recall task precisely because it did seem so difficult. Furthermore, Nairne et al.'s delayed-recall performance seemed unusually high for a Brown-Peterson type of task with eight items per trial (above 60% for the simply read items). Therefore, it was hypothesized that the delayed-recall advantage for the simply read items obtained by Nairne et al. was the result of the odd-even distractor task, whereas the delayed-recall advantage for the generated items obtained by Burns was due to the number-recall task.

## EXPERIMENTS 1A AND 1B

The purpose of Experiments 1A and 1B was to determine whether Nairne et al.'s (1991) delayed-recall results were due to the use of a distractor task that was different from that used by Burns (1986). Experiment 1A was a direct replication of Burns's (1986) Experiment 4. Experiment 1B was identical to Experiment 1A, with the exception that the odd-even distractor task replaced the

number-recall task. If the distractor task was the cause of the difference between the two studies, then Experiment 1A should produce the same results as Burns's experiment, while Experiment 1B should produce delayed-recall results similar to those of Nairne et al.

Subjects were given 12 lists of six items each. Half the lists were read items and the other half were generate items (word fragments with one letter missing). Three retention intervals were given, 0, 40, or 80 sec, and for the two longer retention intervals, either the odd-even or the number-recall distractor task was employed.

### Method

**Subjects.** The subjects were 48 college students enrolled in various psychology courses. They voluntarily participated in return for extra credit toward their final grade. The 24 subjects in each experiment were tested at different times. For this reason, each experiment was analyzed separately.

**Lists and Design.** Twelve lists of six words each were constructed from the norms of Snodgrass and Vanderwart (1980). Words were randomly assigned to lists with the restriction that no obviously strong relationship existed between any two words in a list. The generate items were created by replacing one letter from each word with a hyphen. Each generate item had only one possible solution. Pilot testing was performed to ensure that the subjects could almost always correctly generate the items within 600 msec.

One variable manipulated in Experiments 1A and 1B was whether the items in a list were simply read or generated items. Although this factor was manipulated within subjects, it nonetheless was a between-lists design. The second variable manipulated, retention interval (0, 40, or 80 sec), was varied within subjects. The retention intervals of 40 and 80 sec were filled with distractor activity. Subjects given the number-recall distractor task (Experiment 1A) were required to study and recall one or two lists of six two-digit numbers, respectively, for the 40- and 80-sec-delay conditions. Subjects given the odd-even distractor task (Experiment 1B) were presented with a single digit from 1 to 9 every 1 sec, and they had to determine whether each digit was odd or even. They responded by pressing one of two buttons on the computer keyboard.

Counterbalancing ensured that each list was presented in read or generate form and was tested at each retention interval equally often. The order in which the three different retention intervals was given was held constant. The order was the same for all subjects (i.e., 80-40-0-0-40-80-40-0-80-80-0-40). This order was chosen so that the subjects could not easily guess which retention interval would be given on a particular trial. Experiments 1A and 1B were 2 (stimulus type: read or generate)  $\times$  3 (retention interval: 0, 40, or 80 sec) within-subjects factorial designs.

**Procedure.** The subjects were tested individually in small rooms. The target items were presented by computer one at a time in uppercase letters, and were centered on the monitor screen. For each list, the appearance of the first word was signaled by the word "READY" at the top of the video screen 3 sec prior to display of the target items. Each of the six words in the list was presented for 500 msec, with an interstimulus interval of 100 msec.

Six hundred milliseconds after the offset of the last word in the list, a row of either seven question marks (signaling recall) or seven number signs (signaling the distractor task) appeared in the middle of the screen. For the immediate-recall trials, the row of question marks signaled the subjects to begin recalling the list of words. They were given 60 sec for written recall of the six words on the list. Forty seconds into the recall period, the subjects heard a short beep, and were shown instructions that informed them to finish writing six items. Recall was forced, in that the subjects were required to write six items even if they had to guess. This forced-choice pro-

cedure was used to equate subjects' recall (response) criterion across stimulus type and retention interval.

For the delayed trials, the row of number signs at the end of list presentation signaled the beginning of a 40-sec distractor task. On these trials, the subjects in Experiment 1A were required to study and recall a list of six two-digit numbers during the 40-sec retention interval. The row of number signs was presented for 1,000 msec, followed by six two-digit numbers presented at a 1,350-msec rate. Immediately following the last two-digit number, instructions appeared on the screen informing the subjects to recall the numbers. The subjects were given 30 sec for written forced recall. When the 30-sec recall period was over, the subjects were presented either the number signs again (signaling a second 40-sec distractor period with a new list of numbers) or the row of question marks, signaling recall of the last list of words that was presented. All the subjects were instructed that they were to try to do their best on both the word lists and the number lists.

The subjects in Experiment 1B were given the odd-even digit distractor task on delayed-recall trials. Single digits were displayed on the screen for 1,000 msec each. The subjects were required to press one key if the digit was even and an adjacent key if the digit was odd. The distractor task lasted for 0 (i.e., was not given), 40, or 80 sec, and it was identical to that used by Nairne et al. (1991). When the distractor task was finished, the row of question marks signaling recall of the last list of words was presented.

The procedures for the delayed-recall tests were identical to those employed with the immediate-test trials. Seven seconds intervened between the end of the recall period for each word list and onset of the "READY" signal for the next list of words.

Additionally, all the subjects were told to name each word aloud as it was being presented, and to generate the items when necessary. Generation failures were recorded by the experimenter.

## Results and Discussion

The subjects correctly generated 94.6% of the items in Experiment 1A and 98.5% of the items in Experiment 1B. The recall proportions presented in Table 1 are conditionalized on whether subjects correctly generated the items during input.

A 2 (stimulus type: read or generate)  $\times$  3 (retention interval: 0, 40, or 80 sec) within-subjects analysis of variance (ANOVA) was performed on both the unconditional and the conditionalized recall proportions for each experiment. The conditional and unconditional scores produced the same pattern of results, and so only the conditionalized scores are presented. The ANOVA for Experiment 1A, which employed the number-recall task, revealed that both the main effect of retention interval and the interaction

between stimulus type and retention interval were significant [ $F(2,46) = 102.48$ ,  $MS_e = 0.02$ , and  $F(2,46) = 7.19$ ,  $MS_e = 0.01$ ]. For Experiment 1B (the odd-even distractor), the simply read items were recalled better than the generated items [ $F(1,46) = 4.31$ ,  $MS_e = 0.02$ ], and the main effect of retention interval was also significant [ $F(2,46) = 14.00$ ,  $MS_e = 0.02$ ].

Individual comparisons between the simply read and generated items were performed on both the immediate-recall and the 80-sec delayed-recall proportions for Experiment 1A. The immediate-recall comparison revealed that the simply read items were recalled significantly better than the generated items [ $F(1,23) = 7.43$ ,  $MS_e = 0.01$ ]. However, the 80-sec delayed-recall comparison revealed that the generated items were recalled significantly better than the simply read items [ $F(1,23) = 6.71$ ,  $MS_e = 0.01$ ].

Included in Table 1 is an index of amount of order processing that the subjects performed on each list type. The index, developed by Asch and Ebenholtz (1962), measures the number of times adjacently recalled items preserve the relative order of input, with chance performance at .50 (or 50%). This index is the same one that was used by Nairne et al. (1991). In those few instances in which a score on a particular trial could not be calculated because less than two items were recalled, subjects were given the score they received on the other trial with the same retention interval. For example, if a subject had an order score of .60 on the first generate trial at the 80-sec retention interval and recalled only one item on the second generate trial at the 80-sec interval, then that subject received a score of .60.

The 2  $\times$  3 ANOVAs performed on the input order scores for each experiment revealed that the only significant effect for each experiment was the main effect of stimulus type [ $F(1,23) = 4.95$ ,  $MS_e = 0.08$ , and  $F(1,23) = 8.96$ ,  $MS_e = 0.04$ , respectively, for Experiments 1A and 1B]. No other effects approached significance (largest other  $F = 1.84$ ).

A recall advantage for the simply read items always was obtained in immediate recall, but was obtained in delayed recall only when the presumably easy odd-even distractor task was used. When the number-recall distractor task was employed, an 80-sec delayed-recall advantage was obtained for the generated items. The odd-even distractor

Table 1  
Mean Proportion of Items Recalled and Mean Input-Output Order Scores for the Conditions of Experiments 1A and 1B

Performance Measures	Retention Interval					
	0 sec		40 sec		80 sec	
	Read	Generate	Read	Generate	Read	Generate
Experiment 1A						
Recall scores	.83	.75	.49	.49	.39	.47
Order scores	.74	.67	.63	.58	.73	.52
Experiment 1B						
Recall scores	.76	.70	.64	.59	.61	.58
Order scores	.70	.60	.64	.59	.72	.56

task produced results consistent with those of Nairne et al. (1991), whereas the number-recall distractor task produced results similar to Burns (1986).

The results of Experiments 1A and 1B lead to the suggestion that simply reading the items enhanced serial order processing, relative to generating the items. The additional serial order information for the simply read items enhanced immediate recall, but only enhanced delayed recall when the odd-even distractor task was employed. A discussion of why serial order information improved recall after the odd-even distractor task but not after the number-recall task will be presented following presentation of Experiment 2, which is a conceptual replication of Experiments 1A and 1B.

## EXPERIMENT 2

Experiment 2 served several purposes. First, it was implied in the introduction that the number-recall task was a more difficult distractor task than the odd-even task, and the delayed-recall scores certainly support this contention. However, Experiments 1A and 1B do not directly test whether the delayed-recall performance differences are due to the difficulty level of the distractor. It is possible that some other difference between the two tasks caused the performance differences. For example, it is possible that both tasks are equally difficult (i.e., require the same amount of cognitive resources) but that being required to recall lists of numbers *interferes* more with recalling lists of words than does the odd-even task. Therefore, the difficulty level of the distractor task was directly manipulated in Experiment 2.

The second reason for conducting Experiment 2 was that Experiments 1A and 1B relied on an indirect measure of order processing. Specifically, amount of order information processed was measured by the amount of input-output order correspondence in the free-recall protocols. Nairne et al. (1991) used the more direct order-reconstruction test to measure order processing. Experiment 2 followed the lead of Nairne et al. and implemented an order-reconstruction test.

Finally, an implicit assumption of Experiments 1A and 1B was that although simply reading the items produced more order processing, generating the items produced more processing of item-specific information. In fact, Nairne et al. (1991) showed that performance on an unexpected final recognition test, presumed to be sensitive to item-specific information, was higher for the generated items than for the simply read items. A final unexpected recognition test was given to the subjects in Experiment 2 to verify that the generated items received more item-specific processing.

### Method

**Subjects.** The 72 new subjects participating in Experiment 2 were taken from the same pool as the subjects in Experiments 1A and 1B.

**Design and Procedure.** The same list of 12 six-item lists used in Experiments 1A and 1B were employed in Experiment 2. Half

the lists were tested with a free-recall test (rather than the forced test used in Experiments 1A and 1B), and half were tested with an order-reconstruction test. The reconstruction test consisted of simultaneously presenting all six words on the computer monitor in a random order, and requiring subjects to write the words in the order in which they were originally presented. The subjects were given 60 sec to perform this task.

The distractor task used in Experiment 2 was a number-naming task. The subjects were presented with a four-digit number in the center of the screen and were required to name each digit of the number aloud. In the easy distractor condition, the numbers were presented at a 2-sec rate, whereas for the difficult distractor task the rate was 1 sec per number. Thus, the same general distractor task was given for both the easy and difficult distractor conditions. The only difference was in the speed in which the numbers were presented. The two levels of distractor difficulty should vary only in terms of difficulty level, and not on any other dimension such as interference.

The design of the experiment was a 2 (distractor task: easy or difficult)  $\times$  2 (stimulus type: read or generate)  $\times$  2 (test type: free recall or order reconstruction)  $\times$  3 (retention interval: 0, 40, or 80 sec) factorial, with the last three factors manipulated within subjects. The subjects contributed one trial of data for each of the 12 within-subject cells.

Counterbalancing was more complete in Experiment 2 than it was in the previous experiments. Specifically, the 12 lists of words were presented in the same order for all 12 subjects. However, each list was presented in read or generate form, was tested by free recall or order reconstruction, and was tested at each retention interval equally often. This counterbalancing scheme is superior to the one used in Experiments 1A and 1B because it eliminates the possibility of differential buildup of proactive interference across retention intervals.

Following testing of the final list of words, the subjects were given an unexpected three-alternative forced-choice recognition test. The test contained all 72 target items and 144 lures, which were selected from the same source as the target items. The subjects were given 5 min to complete the test. All other procedures were identical to Experiments 1A and 1B.

### Results and Discussion

The subjects correctly generated 93.7% and 94.3% of the items during input in the difficult and easy distractor conditions, respectively. A 2 (distractor task: easy or difficult)  $\times$  2 (stimulus type: read or generate)  $\times$  3 (retention interval: 0, 40, or 80 sec) mixed-factor ANOVA performed on the conditionalized recall proportions, presented in Table 2, revealed that the main effect of retention interval was significant [ $F(2,140) = 143.32$ ,  $MS_e = 0.04$ ]. In addition, the distractor task  $\times$  stimulus type interaction was significant [ $F(1,70) = 7.64$ ,  $MS_e = 0.03$ ], and the three-way interaction approached significance [ $F(2,140) = 2.82$ ,  $MS_e = 0.03$ ,  $p = .06$ ].

A similar ANOVA performed on the unconditional recall scores produced slightly different results in that the main effect of stimulus type was significant [ $F(1,70) = 5.41$ ,  $MS_e = 0.03$ ], and the interaction between stimulus type and retention interval was significant [ $F(2,140) = 3.20$ ,  $MS_e = 0.03$ ]. As in the first analysis, there was a significant main effect of retention interval and a significant distractor task  $\times$  stimulus type interaction.

Separate 2 (stimulus type)  $\times$  3 (retention interval) within-subjects ANOVAs were performed on the condi-

Table 2  
Mean Retention Scores for the Conditions of Experiment 2

Performance Measures	Retention Interval					
	0 sec		40 sec		80 sec	
	Read	Generate	Read	Generate	Read	Generate
Difficult Distractor Task						
Recall scores	.74	.67	.33	.40	.26	.36
Reconstruction scores	.61	.64	.38	.35	.42	.33
Order scores	.81	.81	.64	.59	.64	.58
Recognition scores	.81	.83	.72	.76	.74	.76
Easy Distractor Task						
Recall scores	.73	.68	.46	.39	.36	.29
Reconstruction scores	.63	.58	.45	.36	.39	.32
Order scores	.77	.78	.67	.59	.66	.58
Recognition scores	.82	.81	.81	.81	.75	.81

tionalized recall scores for both the easy and difficult distractor task groups to explore the significant two-way interactions and the near-significant three-way interaction. The ANOVA for the easy distractor task showed that the simply read items were recalled better than the generated items [ $F(1,35) = 10.20$ ,  $MS_e = 0.04$ ]. The main effect of retention interval also was significant [ $F(2,70) = 62.28$ ,  $MS_e = 0.04$ ], but the interaction did not approach significance ( $F < 1$ ). The ANOVA for the difficult distractor task showed that both the main effect of retention interval and the interaction were significant [ $F(2,70) = 75.54$ ,  $MS_e = 0.04$ , and  $F(2,70) = 3.54$ ,  $MS_e = 0.04$ , respectively].

Individual comparisons between the simply read and generated items at each retention interval were performed on the recall scores for the difficult distractor task group. These comparisons revealed that the simply read items were recalled better than the generated items in immediate recall [ $F(1,35) = 10.22$ ,  $MS_e = 0.02$ ], but that there was no difference at the 40- or 80-sec retention intervals [ $F(1,35) = 1.37$ ,  $MS_e = 0.04$ , and  $F(1,35) = 2.52$ ,  $MS_e = 0.04$ ]. When the scores from the 40- and 80-sec-interval groups were combined, however, the recall advantage for the generated items approached significance [ $F(1,35) = 3.54$ ,  $MS_e = 0.04$ ,  $p = .06$ ]. Similar comparisons performed on the unconditional recall scores revealed that the advantage for the generated items at the 80-sec retention interval was significant [ $F(1,35) = 4.36$ ,  $MS_e = 0.04$ ], as was the combined 40- and 80-sec retention interval comparison [ $F(1,35) = 5.07$ ,  $MS_e = 0.02$ ].

The reconstruction test results were analyzed in two ways. First, we used the Asch-Ebenholtz procedure employed in Experiments 1A and 1B. Second, we calculated the probability that an item was placed in the correct serial position in which it was originally presented. Both the Asch-Ebenholtz order scores and the reconstruction scores are displayed in Table 2. A 2 (distractor task)  $\times$  2 (stimulus type)  $\times$  3 (retention interval) ANOVA was performed on the scores derived from both scoring procedures. Both scoring procedures produced a significant main effect of retention interval [smallest  $F(2,140) = 31.73$ ,  $MS_e = 0.05$ ], but only the Asch-Ebenholtz score

produced a significant advantage for the simply read items [ $F(1,70) = 6.51$ ,  $MS_e = 0.03$ ]. The advantage for the simply read items in terms of the probability of correct serial placement approached significance [ $F(1,70) = 3.46$ ,  $MS_e = 0.07$ ,  $p = .07$ ].

The unconditional final unexpected recognition scores are presented in Table 2. A 2 (distractor task)  $\times$  2 (stimulus type)  $\times$  2 (type of previous test: recall or reconstruction)  $\times$  3 (retention interval) ANOVA performed on the recognition probabilities showed that the generated items were recognized better than the simply read items [ $F(1,70) = 5.73$ ,  $MS_e = 0.02$ ], and the previously reconstructed lists were recognized better than the previously recalled lists [ $F(1,70) = 168.56$ ,  $MS_e = 0.02$ ]. In addition, retention interval interacted both with distractor task and type of test [ $F(2,140) = 3.79$ ,  $MS_e = 0.02$ , and  $F(2,140) = 3.45$ ,  $MS_e = 0.02$ , respectively].

Individual comparisons performed on the recognition scores revealed that the significant retention interval  $\times$  distractor task interaction resulted from recognition being higher for the easy distractor task condition, but only at the 40-sec retention interval [ $F(1,70) = 4.75$ ,  $MS_e = 0.01$ ]. Similarly, the retention interval  $\times$  type of test interaction was due to the fact that performance decreased from 0 to 80 sec for the previously recalled items [ $F(2,142) = 4.65$ ,  $MS_e = 1.83$ ], but not for the previously reconstructed items.

The free-recall results of Experiment 2 generally replicate those of Experiments 1A and 1B. The simply read items were recalled better than the generated items when an easy distractor task was used. However, when a difficult distractor task was used, the simply read items were recalled better than the generated items in immediate recall only. The generated items were actually recalled better in delayed recall (at least for the unconditional recall scores). Experiment 2 supports our contention that the critical difference between the distractor tasks used in Experiments 1A and 1B is difficulty level: The number-recall task was more difficult than the odd-even task.

The overall advantage for the simply read items on the reconstruction test (as shown by the Asch-Ebenholtz scoring procedure) suggests that the simply read items were

given more serial order processing than the generated items. However, the enhanced serial order processing only facilitated recall when testing was immediate or when an easy distractor task was given. The order information was not very beneficial to long-term recall when a difficult distractor task was employed. The main finding concerning the recognition scores is that recognition was higher for the generated items than for the simply read items, suggesting that the generated items received more item-specific processing.

### GENERAL DISCUSSION

Burns (1986, Experiment 4) showed that generated items were recalled better than simply read items in delayed recall. Using similar procedures, Nairne et al. (1991) observed a delayed-recall advantage for the simply read items. The purpose of Experiments 1A and 1B was to determine if the difference in distractor tasks used by Burns and Nairne et al. was responsible for the difference in the results. Experiment 1A was a direct replication of Burns's experiment, and Experiment 1B was a direct replication of Experiment 1A, with the exception that the distractor task used by Nairne et al. replaced the number-recall task. The results of Experiment 1A closely replicated those of Burns (1986), and Experiment 1B produced results consistent with Nairne et al.'s results.

In Experiment 2, difficulty level of the distractor task was treated as an independent variable, and both an order-reconstruction test and a final unexpected recognition test were given. The free-recall scores produced the same pattern of results as those obtained in Experiments 1A and 1B: Higher free recall for the simply read items in immediate recall and in delayed recall when an easy distractor task was given, but higher delayed recall for the generated items when a difficult distractor task was used. It was concluded that difficulty level was the important difference between the distractor tasks used in Experiments 1A and 1B.

The reconstruction scores generally showed that simply reading the items resulted in greater processing of order information than generating them, regardless of retention interval. Finally, the recognition scores showed that generated items were more accurately recognized than simply read items, regardless of retention interval, test type, or distractor difficulty.

The overall pattern of results suggests that generating inhibits the processing of serial order information (as measured by the reconstruction test), but enhances the processing of item-specific information (as measured by the recognition test). This conclusion is consistent with previous research (e.g., Burns, 1992; Nairne et al., 1991; Serra & Nairne, 1993). It also appears that the additional serial order information given to the simply read items enhances immediate free-recall performance. However, the serial order information enhances delayed-recall performance only when an easy distractor task is used. When a difficult distractor task is used, any memorial benefits

of serial order information that exist appear to be completely masked by the enhanced processing of item-specific information given to the generated items.

It is well known that rehearsal can occur during distractor activity, and the easier the distractor activity the greater the opportunity for rehearsal (e.g., Kroll & Kellicutt, 1972). Therefore, it seems likely that the easy distractor task allowed the subjects to rehearse the target items during the distractor activity and this rehearsal protected them from forgetting. The recall proportions showed a much greater decrease across the retention intervals for the difficult distractor task than for the easy task, supporting the contention that the subjects were rehearsing during the easy distractor task. In addition, upon questioning, several students in the easy distractor task conditions of Experiments 1B and 2 stated that they often rehearsed during the distractor task, whereas very few students in the difficult distractor task conditions of Experiments 1A and 2 reported rehearsing during the distractor period.

It is possible that the additional rehearsal given to the subjects in the easy distractor conditions decreased forgetting of the order information. This would have made the order information more accessible during delayed recall. Hence, the additional order information given to the simply read items relative to the generated items could have facilitated recall.

The fact that retention interval did not interact with difficulty level of the distractor task on the reconstruction test is inconsistent with the notion that additional rehearsal during the easy distractor task decreased forgetting of the order information. Therefore, there is little support for the view that the difficult distractor task caused greater forgetting of the order information.

A second explanation is that the rehearsal allowed many of the items to be maintained in short-term memory, thus making the long-term recall test functionally a short-term-memory test for those items. The high delayed-recall scores obtained by the subjects given the easy distractor task (especially in Experiment 1B) are consistent with the view that the subjects were maintaining items in short-term memory. At this point, however, the hypothesis that contributions from short-term memory influenced the delayed-recall scores is purely speculative.

Nairne et al. (1991) and Serra and Nairne (1993) proposed the item-order hypothesis for explaining design controversies in the generation effect literature. The hypothesis is similar to the multifactor theories of the generation effect that have been proposed (e.g., Burns, 1990, 1992; Hirshman & Bjork, 1988; McDaniel, Waddill, & Einstein, 1988). The multifactor theories assume that generating influences the degree to which item-specific and relational processing are performed.

The item-order hypothesis can be seen as a variant of the multifactor theories, in that order information can be viewed as a specific type of relational information (see Burns, 1992). The present results, however, question the viability of the item-order hypothesis in accounting for the negative generation effect obtained with the more

traditional long-term recall tasks (e.g., Burns, 1992; Schmidt & Cherry, 1989). This is because the additional order information does not always appear to facilitate free recall. In fact, the simply read items, which received more order information, sometimes produced worse recall than the generated items. Clearly, more research is needed to determine if, and/or when, order information facilitates long-term recall performance.

Both Burns (1992) and Nairne et al. (1991) concluded that much of the research on the generation effect could be explained by the processing tradeoff between individual-item information and order information. Generation enhances processing of individual-item information, but hinders processing of order information. The present experiments support the notion of a tradeoff between item and order information, but they do not provide support for the notion that this tradeoff is responsible for the lack of generation effects (and the occasional recall advantage for the simply read items) in between-lists designs.

It seems that additional research is needed before one speculates on the relative contribution of order information in producing the complex set of results found in the generation effect literature. A similar conclusion can be drawn concerning the role of the item-order distinction in long-term retention.

#### REFERENCES

- ASCH, S. E., & EBENHOLTZ, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *Journal of Psychology*, *54*, 3-31.
- BEGG, I., & SNIDER, A. (1987). The generation effect: Evidence of generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *13*, 553-563.
- BURNS, D. J. (1986). *Differential rates of forgetting: The exception or the rule?* Unpublished doctoral dissertation, State University of New York at Binghamton.
- BURNS, D. J. (1990). The generation effect: A test between single- and multifactor theories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 1060-1067.
- BURNS, D. J. (1992). The consequences of generation. *Journal of Memory & Language*, *31*, 615-633.
- EINSTEIN, G. O., MCDANIEL, M. A., BOWERS, C. A., & STEVENS, D. T. (1984). Memory for prose: The influence of relational and proposition-specific processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 133-143.
- HIRSHMAN, E., & BJORK, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 484-494.
- KROLL, N. E., & KELLICUTT, M. H. (1972). Short-term recall as a function of covert rehearsal and of intervening task. *Journal of Verbal Learning & Verbal Behavior*, *11*, 196-204.
- MCDANIEL, M. A., WADDILL, P. J., & EINSTEIN, G. O. (1988). A contextual account of the generation effect: A three factor theory. *Journal of Memory & Language*, *27*, 521-536.
- NAIRNE, J. S. (1990). Similarity and long-term memory for order. *Journal of Memory & Language*, *29*, 733-746.
- NAIRNE, J. S., RIEGLER, G. L., & SERRA, M. (1991). Dissociative effects of generation on item and order information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 702-709.
- SCHMIDT, S. R., & CHERRY, K. (1989). The negative generation effect: Delineation of a phenomenon. *Memory & Cognition*, *17*, 359-369.
- SERRA, M., & NAIRNE, J. S. (1993). Design controversies and the generation effect: Support for an item-order hypothesis. *Memory & Cognition*, *21*, 34-40.
- SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 592-604.
- SLAMECKA, N. J., & KATSAITI, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory & Language*, *26*, 589-607.
- SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 174-215.

(Manuscript received June 11, 1992;  
revision accepted for publication May 18, 1993.)