

The negative generation effect: Delineation of a phenomenon

STEPHEN R. SCHMIDT and KATIE CHERRY

Virginia Polytechnic Institute and State University, Blacksburg, Virginia

In four experiments employing between-list designs, generation was found to have negative effects on free recall of word pairs and on cued recall of the second word. In addition, generation had negative effects on measures of word-pair integration and on clustering in recall. In contrast, positive effects of generation were found on free recall of second words alone, and on a recognition test for memory of the second word. It was concluded that in between-list designs, generation led to greater individual-item processing of the generated term than reading, but this processing occurred at the expense of processing the relation between the words in a pair and processing the relations between different pairs in a list.

Slamecka and Graf (1978) demonstrated that recognition of the second word from word pairs was greater if subjects were required to fill in missing letters from that word than if subjects were simply asked to read the word pairs. This phenomenon has been called the generation effect. The generation effect has been demonstrated with word pairs (Slamecka & Graf, 1978), single words (Glisky & Rabinowitz, 1985), sentences (Graf, 1982), and prose materials (McDaniel, 1984). In addition, the generation effect has been found with cued recall and recognition tests (Slamecka & Graf, 1978), and under both intentional (Slamecka & Graf, 1978) and incidental (Glisky & Rabinowitz, 1985; Jacoby, 1978; McDaniel, 1984) instructions.

The generation effect is important for both its practical and its theoretical implications. From a practical point of view, leaving letters out of textual material may be an easy way to improve memory of a text. From a theoretical point of view, the generation effect potentially provides evidence relevant to a number of theoretical frameworks. For example, perhaps the generation of material requires greater depth of processing (Craik & Lockhart, 1972) than reading. Perhaps generation should be thought of as a more difficult encoding task than reading, and thus the effects of generation should be cast within an encoding-difficulty framework (e.g., Tyler, Hertel, McCallum, & Ellis, 1979). Alternatively, the generation effect can be approached from the perspective of the distinctiveness hypothesis and the distinction between individual-item and relational processing (Einstein & Hunt, 1980; Hunt & Einstein, 1981).

Experiment 1 reported below was designed to determine the effects of generation on individual-item and relational processing. *Individual-item* processing refers to processing the specific attributes of each to-be-remembered item. *Relational processing* refers to processing the relations between to-be-remembered events. Three hypotheses concerning the effects of generation on individual-item and relational processing have been proposed: (1) the specific enhancement hypothesis, in which generation enhances the processing of the relation between the first term and the generated term in a word pair (Rabinowitz & Craik, 1986), (2) the "individual-item" hypothesis, in which generation enhances individual-item information of the generated word (McDaniel, Einstein, Dunay, & Cobb, 1986), and (3) the "more of both" hypothesis, in which generation leads to enhanced relational processing of a pair and to enhanced processing of the response term (Hirshman & Bjork, 1988).

The original intent of this research project was to distinguish among these hypotheses. However, this intent was sabotaged by serendipity, for we employed between-subjects designs. Recent evidence suggests that the generation effect is obtained only in within-list designs (Begg & Snider, 1987; Slamecka & Katsaiti, 1987). In a within-list design, subjects read some of the items on a to-be-remembered list and generate other items on the same list. Memory of the generated items is then compared with memory of read items from the same list. In such a design, generated items may receive extra attention at the expense of read items (Slamecka & Katsaiti, 1987). In between-list designs, subjects either read or generate the entire list, and comparisons are made between memory of the read and memory of the generated lists. Slamecka and Katsaiti (1987) demonstrated that positive effects of generation are found on the free recall of the generated terms in within- but not in between-list designs. In addition, items read were more poorly recalled in the within-list design than in the between-list design, suggesting that the enhanced recall of generated items in the within-list

We extend our special thanks to Ronald Davis, Kevin Snow, and Maureen Bergen for their assistance in data collection and scoring. We also wish to thank Constance R. Schmidt and Jan C. Rabinowitz for their comments on an earlier version of the manuscript. Experiment 3 was presented at the Annual Meeting of the Midwestern Psychology Association, May 8, 1987. Requests for reprints should be sent to Stephen R. Schmidt, Department of Psychology, Middle Tennessee State University, Murfreesboro, TN 37132.

design occurred at the expense of the read items. In light of Slamecka and Katsaiti's (1987) results, as well as the results of Begg and Snider (1987), the three hypotheses described above are in need of reevaluation. Generating words may indeed influence individual-item and/or relational processing, but that influence may depend upon list structure.

In the experiments reported below, we employed between-list manipulations of encoding tasks (read vs. generate). To our surprise (prior to the publication of Begg & Snider, 1987, and Slamecka & Katsaiti, 1987), we found that generation led to *lower* levels of free recall than reading. We then pursued this "negative generation effect," finding it under a variety of conditions across seven different experiments involving more than 1,000 subjects. Five of these experiments are reported below. In these experiments, we explored the negative generation effect from the perspective of individual-item and relational processing. Again, several hypotheses can be generated. In a between-list design, it is unlikely that generation has a positive effect on both relational and individual-item processing and yet produces a negative effect on recall. However, it is possible that generation has a negative effect on one type of processing and not the other when compared to reading. Or perhaps, generation may increase one type of processing while decreasing the other.

Positive effects of generation in between-list designs have been reported, most notably by Slamecka and Graf (1978, Experiment 1) and Graf (1982, Experiment 1). A recognition test of memory was employed in both of these experiments, suggesting that, even in between-list designs, generation enhances the individual-item processing important for recognition judgments. However, individual-item processing is also important in recall. Slamecka and Katsaiti's (1987) failure to obtain a generation effect on recall in between-list designs thus challenges the notion that generation increases individual-item processing in this type of design. Perhaps generation has a positive effect on individual-item processing and a negative effect on relational processing, leading to no effect (and sometimes negative effects) of generation on recall. We explored this possibility in the experiments reported below.

EXPERIMENT 1

As noted above, the original intent of this investigation was to determine the roles of individual-item and relational processing in the generation effect. We varied the strength of the associative relation between words in a pair (strong vs. weak) and type of memory test (cued vs. free recall). According to the specific-enhancement hypothesis, generation enhances the learning of the association between words in an associated pair, and thus increases relational processing. This led us to predict larger positive effects of generation with weakly related pairs than with strongly related pairs. We predicted this effect on both cued and free recall measures of memory.

Based on the individual-item hypothesis, we predicted generation effects of equal magnitude with strongly and weakly related pairs. However, the generation effect should depend on the type of memory test. In cued recall, second-term processing is very important, and thus the individual-item hypothesis predicted a positive effect of generation in cued recall. In the free recall of pairs, pair integration (i.e., relational processing) is of primary importance, and thus the increased processing of the second terms alone should have little effect on the recall of intact pairs. If generation enhances both relational and individual-item processing, then the generation effect should be found with both strong and weak pairs and on both measures of memory. In summary, the relational hypothesis predicted an interaction between task (read vs. generate) and associative strength, but no interaction with type of memory test. The individual-item hypothesis predicted an interaction between task and type of memory test, but no interaction with associative strength. The "more of both" hypothesis predicted general positive effects of generation independent of type of pair and type of memory test.

In addition to addressing the issues discussed above, in Experiment 1 we extended the study of the generation effect by employing two levels of generation difficulty. Gardiner, Smith, Richardson, Burrows, and Williams (1985) reported that the positive effects of generation on cued recall (in a within-list design) increased as the number of letters omitted from the words to be generated increased. If these effects are the result of individual-item processing, relational processing, or both, it should be possible to demonstrate changes in the relevant type of processing by means of changes in the difficulty of the generation task.

Method

Subjects and Design

The participants were 222 undergraduate students enrolled in introductory psychology classes, who received extra credit in return for their participation. These subjects participated in a 3 (encoding difficulty: read vs. easy completion vs. hard completion) \times 2 (type of associative relationship: strong vs. weak) \times 2 (type of recall: free vs. cued) factorial. The associative relationship factor was treated within subjects; all other manipulations were between subjects. The number of subjects contributing to each cell is reported in Table 1.

Materials

Fourteen strongly associated noun pairs and an equal number of weak associates were selected from the Palermo and Jenkins (1964) association norms. The mean associative ratings (combined across male and female college students) were 5.48 and 3.14 for the strong and weak pairs, respectively. Examples of the strong and weak pairs are MAN-WOMAN and MOON-ROCKET, respectively. Note that even with the weak pairs, generation of the second term should be guided by the first term. Thus the specific-enhancement hypothesis predicted positive effects of generation for both the strong and weak pairs. For the *easy* completion condition, an average of 1.3 letters was omitted from the response member, and an average of 2.6 letters was deleted in the *hard* completion task. An equal num-

ber of letters was removed in the *strong* and *weak* associate conditions to ensure a balance between pair-associative-strength values and encoding difficulty. The pairs were randomly ordered for presentation. Booklets were constructed with a cover sheet, a second page on which the word pairs were listed in a single column, a third page containing addition problems, and a recall sheet.

Procedure

In the *generation* conditions, the subjects were told that they would be asked to fill in the missing letters in word pairs. In the *read* conditions, they were asked to read intact pairs. In both conditions, they were instructed to “concentrate on how the words in each pair are conceptually related or associated with each other,” and to try to learn the word pairs. Prior to presentation, the cued recall test was described to the subjects given cued recall. Similarly, the subjects who received free recall were told about the free recall test. The subjects tested for free recall were asked to recall the word pairs, but if they could only remember a single word, they were asked to record that member alone. They were then given 3 min to either read through and learn the noun pairs, or correctly complete the second member and learn the pairs. The subjects were then instructed to turn to the next page of the test package and solve addition problems for 1 min. After the brief distractor task, they turned to the final page of the test package and were given 3 min to complete the memory test.

Results

For the generation conditions, the probability of correct recall of a pair was calculated for correctly completed pairs only. The data from the cued and free recall tests were subjected to separate analyses of variance. Within each recall condition (cued vs. free), mean completion accuracy and memory performance were evaluated by separate ANOVAs. For all the statistical tests reported below, the traditional $p < .05$ was adopted. All the results are reported in terms of the probability of correct responses. A summary of the results is presented in Table 1.

Cued Recall

Completion. The mean proportion of completions decreased as encoding difficulty moved from easy completion ($M = .99$) to hard completion ($M = .89$), a trend confirmed by a main effect for encoding difficulty [$F(1,68) = 87.15, MS_e = .0046$]. Strong associates were more likely to be completed in both the easy and hard

tasks ($M = .97$) than were the weak associates [$M = .81; F(1,68) = 137.52, MS_e = .0017$]. The encoding difficulty \times associative strength interaction was also reliable [$F(1,68) = 122.53, MS_e = .0017$]. The effects of encoding difficulty were larger with weak than with strong associates.

Recall. The effect of encoding difficulty on recall was reliable [$F(2,106) = 3.75, MS_e = .0138$]. Cued recall in the read group ($M = .98$) exceeded cued recall in the easy completion group ($M = .94$), which exceeded recall in the hard completion group ($M = .92$). In addition, the associative strength factor exerted a significant effect on recall [$F(1,106) = 30.95, MS_e = .0085$]. The mean recall rate for the strongly associated noun pairs ($M = .98$) exceeded the rates of recall for the weak associates ($M = .91$). The encoding difficulty \times associative strength interaction was also significant [$F(2,106) = 3.20, MS_e = .0085$]. This interaction may be the result of a ceiling effect in the cued recall of strongly related pairs (see Table 1). To test this hypothesis, the combined cued recall performance was split at the median and the analyses were repeated on the lower half of the subjects. The same pattern of results was found, with a main effect of associative strength [$F(1,52) = 36.42, MS_e = .0127$], a main effect of encoding difficulty [$F(1,52) = 5.57, MS_e = .0163$], and a significant interaction between encoding difficulty and associative strength [$F(2,52) = 3.29, MS_e = .0127$]. Whereas this interaction was predicted by the relational hypothesis, the interaction was of the wrong form to support the hypothesis. The relational hypothesis predicted a larger positive effect of generation with weakly related pairs. The results demonstrated a larger negative effect of generation with weakly than with strongly related pairs. Of course this result is difficult to explain within the individual-item and “more of both” frameworks, which predicted equal and positive effects of generation with both strongly and weakly related pairs.

Free Recall

Completion. For the easy completion group, completion performance was nearly perfect ($M = 1.00$), whereas the hard completion task led to lower probabilities of completion [$M = .91; F(1,73) = 106.86, MS_e = .0029$]. Again, strong associates were more likely to be completed ($M = .98$) than weak associates [$M = .84; F(1,73) = 70.23, MS_e = .0028$]. In addition, the difficulty \times associative strength interaction was significant [$F(1,73) = 62.93, MS_e = .0028$], once again indicating a larger effect of difficulty with weakly than with strongly related pairs.

Recall of pairs. Encoding difficulty exerted a reliable effect on the free recall of pairs [$F(2,110) = 8.28, MS_e = .0227$]. As in cued recall, the highest probabilities of recall were noted in the read group ($M = .46$), followed by the easy completion group ($M = .45$) and finally the hard completion group ($M = .37$). The main effect of associative strength indicated that strong associ-

Table 1
Comparisons of Cued and Free Recall as a Function of Encoding Difficulty and Associative Strength

	Strong			Weak		
	Read	Easy	Hard	Read	Easy	Hard
Cued Recall						
Completion		1.00	.966		.992	.808
Recall	.989	.974	.976	.961	.899	.872
<i>n</i> *	39	36	34	39	36	34
Free Recall						
Completion		1.00	.978		.996	.837
Pairs Recalled	.547	.527	.447	.377	.367	.287
<i>n</i> *	38	40	35	38	40	35

*Strong versus weak associates was a within-subjects factor.

ates ($M = .51$) were more likely to be recalled than weak associates [$M = .35$; $F(1, 110) = 109.78$, $MS_e = .0137$]. The difficulty \times associative strength interaction did not approach significance ($F < 1.0$, $MS_e = .0127$).

Discussion

The most important findings in Experiment 1 were the negative generation effects found in cued and free recall. These negative effects of generation seriously challenged the relational, the individual-item, and the "more of both" hypotheses regarding the generation effect. The fact that the negative generation effect increased with the difficulty of the generation task provided additional legitimacy for the effect. We pursued the negative generation effect in the following experiments.

EXPERIMENT 2

One explanation of the results above is that, when compared to reading, generation increases individual-item processing of the second term at the expense of processing the relation between words in a pair. Perhaps as the generation task becomes more difficult, individual-item processing increases at the expense of relational processing. This explanation is consistent with several previous findings. As noted above, Slamecka and Graf (1978) reported a positive effect of generation on second-term recognition in a between-subjects design. In addition, they found that generation improved the recognition of the second word in the word pair but not recognition of the first word.

In order to investigate this tradeoff between individual and relational processing, we focused on the free recall of word pairs. In this task, one can look at the effects of generation on the recall of intact pairs, first terms alone, and second terms alone. If the positive generation effect is confined to the free recall of the second terms, then the effect is probably due to the increased individual-item processing of the second terms. To determine the effects of generation on relational processing, a 2×2 contingency table was constructed for each subject, on which we recorded the frequency of recalling both the first and the second terms from a pair, the first but not the second terms, the second but not the first terms, and neither term. From this classification, a measure of association (ϕ) was calculated to determine the effect of generation on the dependence between first- and second-term recall. The ϕ statistic is equivalent to the Pearson product-moment coefficient for dichotomous data (recalled vs. not recalled). If, as suggested by Experiment 1, generation has a negative effect on pair integration, this negative effect should be evident in a lower mean ϕ for the generation conditions than for the read conditions.

One criticism of Experiment 1 was that the subjects were specifically asked to note the relation between the first and second words in the word pairs. Subtle manipulations of instructions have been found to influence the

generation effect. However, it seems unreasonable to attribute the negative generation effect to our instructions alone. Consider, for example, the instructions employed by Rabinowitz and Craik (1986). "For intact word pairs, subjects were instructed to study the two words and concentrate on the meaningful relationship between the words" (p. 233). These instructions are nearly identical to the ones employed in Experiment 1, yet Rabinowitz and Craik found a positive effect of generation in their within-list design. Nonetheless, to address this potential criticism we directly manipulated the instructions given to the subjects in Experiment 2. As in Experiment 1, some subjects were asked to associate the first and second words in a related word pair. The performance of these subjects was then compared to that of subjects asked to copy the second word. These processing instructions were crossed with two encoding tasks, read versus generate.

Method

Subjects and Design. In Experiment 2, 154 subjects participated in a 2×2 between-subjects factorial combining encoding difficulty (read vs. generate) and instructions (associate vs. copy). The number of subjects contributing to each cell is reported in Table 2.

Materials. Twenty-eight word pairs were chosen from the Palermo and Jenkins (1964) word association norms. The word pairs were concrete nouns with moderate associative values. The mean associative value was 33.0 for the male and female college students combined. Some examples of the word pairs are: BREAD-WATER, CITY-COUNTRY, and DOCTOR-LAWYER. For the generation condition, an average of 2.21 letters was omitted from the second word. Word pairs were randomly ordered for presentation.

Procedure. The subjects were told that they would either have to fill in the missing letters in the word pairs (*generation* condition) or to read word pairs. In the *associate* conditions, they were further instructed to "concentrate on how the words in each pair are conceptually related or associated with each other." In the *copy* conditions, the instructions were to "copy the second word of the pair into the blank provided." The subjects in all conditions were told that they were to try to learn the word pairs. Following 3 min of acquisition, the subjects were instructed to turn to the next page of the test package and solve addition problems for 1 min. After the brief distractor task, they turned to the final page of the test package and were given 3 min to complete the memory test. The subjects were instructed to try to recall the word pairs. However, the instructions stressed that "if you can remember only one word from a pair, please record that word." These recall instructions were given prior to presentation and then repeated immediately prior to the memory test.

Results

Four dependent variables were calculated: the proportion of pairs correctly recalled, the proportion of first terms recalled alone, the proportion of second terms recalled alone, and ϕ . The probability of correct completion was also scored for the subjects in the generation condition. Pairs were scored as correctly recalled independently of the order in which the subjects recorded the first and second terms. However, the terms had to be recalled on adjacent blanks of the recall sheet to be scored as a correct pair. In the generation condition, recall was only scored on correctly completed pairs, and each of the mea-

asures was conditionalized on correct completion. Each of these dependent measures is discussed below and a summary of the results is presented in Table 2.

Completion. The associate instructions led to a greater probability of completion ($M = .98$) than did the copy instructions [$M = .95$; $F(1,73) = 7.13$, $MS_e = .0031$].

Pairs recalled. The probability of pair recall was greater with associate instructions ($M = .41$) than with copy instructions [$M = .34$; $F(1,150) = 12.52$]. In addition, the read task led to greater pair recall ($M = .40$) than did the generation task [$M = .35$; $F(1,150) = 7.61$]. The effect of instructions did not interact with task [$F(1,150) = 3.09$]. The MS_e for all tests of pair recall was .0146. These results support the hypothesis that the instructions to associate versus to copy the word pairs significantly influenced the encoding processes. However, they do not support the hypothesis that the specific instructions employed in Experiment 1 were responsible for the negative effects of generation. Negative effects of generation were obtained with both the associate and the copy instructions.

First-term recall. The probability of recalling first terms in isolation was not reliably affected by type of instructions [$F(1,150) = .41$]. Whereas the generation task led to greater first-term recall ($M = .017$) than did the read task [$M = .009$; $F(1,150) = 4.30$], the magnitude of this effect was quite small. The interaction between instructions and task did not approach significance ($F < 1.0$). The MS_e for all tests of first-term recall was .0005.

Second-term recall. There was a significant effect of instructions on the probability of recalling second terms in isolation [$F(1,150) = 6.28$]. The copy instructions led to greater second-term recall ($M = .029$) than did the associate instructions ($M = .014$). There was also a significant effect of task [$F(1,150) = 12.88$]. The generation task led to greater recall ($M = .033$) than did the read task ($M = .011$). The interaction between instructions and task did not approach significance ($F < 1.0$). The MS_e for all tests involving second terms was .0014.

Phi. The measure of dependence between first- and second-term recall revealed main effects of instructions [$F(1,150) = 14.29$] and task [$F(1,150) = 9.93$]. The associate instructions ($M = .62$) led to greater pair dependency in recall than did the copy instructions ($M = .54$). Reading led to greater pair dependency ($M = .61$) than did generation ($M = .55$). The interaction between in-

structions and task was not significant [$F(1,150) = 1.58$]. The MS_e for these tests was .0151.

Discussion

The results of Experiment 2 led to three major conclusions. First, Experiment 2 replicated the negative generation effect found in the free recall of word pairs reported in Experiment 1. The negative generation effect was found with high and low associates in Experiment 1, and in Experiment 2 with medium associates and two different types of instructions (associate vs. copy). Second, the negative effects found in pair recall may be partially attributed to a disruption in pair integration as measured by phi. Thus, generation has a negative effect on within-pair relational processing in a between-list design. Third, positive effects of generation were found on the recall of second terms alone, suggesting that generation had a positive effect on individual-item processing of the second term. This result is important because it was obtained in a between-list design revealing overall negative effects of generation. These conclusions were further evaluated in Experiment 3.

EXPERIMENT 3

One source of the negative generation effect appears to be a decrease in the processing of the association between members of a word pair. However, generation may also lead to a decrease in the processing between pairs in a list. This between-pair relational processing should be important for free recall, and it may be partially responsible for the negative effects of generation observed above. This hypothesis was evaluated in Experiment 3. A categorized word list was employed so that measures of organization in recall could be made. We predicted that generation would lead to lower levels of organization in recall than would reading the material.

A second purpose of Experiment 3 was to extend the findings of the above experiments to a single-pair mode of presentation. In the first two experiments, the word lists were simultaneously presented. The subjects in the read conditions may have scanned up and down the list while studying the material. In the generation condition, there may have been less time for such up and down processing, because the subjects were occupied with generating the to-be-remembered items. Thus it was possible that the negative generation effect obtained in the earlier experiments might not be obtained with single-pair presentation. In Experiment 3, mode of presentation, single-pair versus simultaneous, was combined factorially with the encoding difficulty manipulation employed in Experiment 1.

Method

Subjects and Design. A total of 235 subjects served in 6 different experimental conditions. Three levels of difficulty of encoding (read, easy completion, and hard completion) were combined with

Table 2
Summary of Results from Experiment 2 as a
Function of Task and Processing Instructions

	Associate		Copy	
	Read	Generate	Read	Generate
Pairs recalled	.453	.365	.349	.330
First terms	.011	.018	.008	.016
Second terms	.005	.024	.017	.041
Phi	.659	.572	.560	.522
<i>n</i>	40	38	39	37

two modes of presentation (simultaneous vs. single-pair) in a between-subjects design. The number of subjects contributing to each cell is reported in Table 3.

Materials. Seven categories were selected from the Battig and Montague (1969) category norms. The categories were: BODY PARTS, FRUIT, SPICES, EARTH FORMATIONS, KITCHEN UTENSILS, ARTICLES OF CLOTHING, and VEGETABLES. Eight items were selected from each category. After excluding the four most frequent category exemplars, the next eight exemplars were selected on the basis of two criteria: (1) they were unambiguous members of the category, and (2) they could easily be assigned to pairs. For example, the following pairs were selected from the category BODY PARTS: HANDS-FEET, KNEES-ELBOWS, WRIST-ANKLE, and TOES-FINGERS. One item in each pair was randomly assigned the first position in the pair, while the second member served as the second item in each pair. Three different versions of the pairs were then constructed: read, easy completion, and hard completion. In the easy completion condition, an average of 1.21 letters was replaced with underlined spaces. In the hard completion condition, 2.32 letters on the average were replaced with spaces.

The 28 pairs constructed for each condition were randomly ordered for presentation. The same random order was used for all conditions and subjects. Two types of booklets were constructed. For the simultaneous condition, the 28 pairs were listed in a single column on one page in a booklet. For single-pair presentation, booklets were constructed with one pair on each page of the booklet. All the booklets contained a cover sheet on the front, and an arithmetic sheet followed by a recall sheet at the back.

Procedure. With one exception, the procedure for this experiment was identical to the procedure used in Experiment 2. In the single-pair presentation condition, the subjects were told to turn pages in their booklets to read or generate the word pairs. The instructions stressed that subjects were to proceed through their booklets in a front-to-back fashion, and that they were not to look back at a page once they went on to the next page. Two subjects failed to follow these instructions; 1 subject in the easy completion task and 5 subjects in the hard completion tasks did not finish the task in the allotted time. The data from these subjects were excluded from all analyses. As in Experiment 2, the recall instructions encouraged the subjects to report partially recalled pairs.

Results

The four dependent measures employed in Experiment 2 were again calculated. The phi measure of dependence between first- and second-term recall provided an index of within-pair relational processing. In order to investigate the effects of generation on between-pair relational processing, two measures of organization in recall were calculated: the number of categories recalled and the ratio of repetition (Bousfield, 1953). Hunt and Seta (1984) have argued that increases in relational processing are associated with increases in category recall and the ratio-of-repetition measure of clustering. Because we were interested in recall of word pairs as a result of between-pair processing, recall of partial pairs was ignored in these calculations. Thus the ratio of repetition was the number of category repetitions (recall of pairs from the same category on adjacent lines), divided by the number of pairs recalled minus one. Partial pairs were not counted as repetitions, but they were counted as interruptions if the recalled part was from a different category than the preceding pair was. A summary of

the results for these six dependent measures is reported in Table 3. Separate ANOVAs were calculated for each measure.

Completion. Successful completion was greater for the easy completion task ($M = .99$), than for the hard completion task [$M = .97$; $F(1,142) = 9.91$]. Successful completion was also greater in the simultaneous mode of presentation ($M = .99$) than in the single-pair mode [$M = .97$; $F(1,142) = 6.94$, $MS_e = .0025$ for both tests]. No other reliable effects were found in the completion data.

Pairs recalled. The recall of intact pairs declined as encoding difficulty increased [$F(2,221) = 7.43$]. For read, easy completion, and hard completion, the mean recall levels were .472, .425, and .395, respectively. Recall was greater following simultaneous presentation ($M = .482$) than following single-pair presentation [$M = .382$; $F(1,221) = 35.99$]. The interaction between mode of presentation and difficulty was not reliable [$F(2,221) = 2.08$]. The MS_e for all tests was .0159. These results replicated the findings of the previous experiments, in that pair recall declined as encoding difficulty increased. The results extended these findings to single-pair presentation and to the recall of categorized word lists.

First-term recall. An analysis of first-term recall indicated little effect of difficulty with mean recall probabilities of .030, .029, and .038 for the read, easy completion, and hard completion groups, respectively [$F(2,221) = .76$]. There was no effect of mode of presentation, and mode of presentation did not interact with difficulty ($F_s < 1.0$). The MS_e for all tests was .0018.

Second-term recall. For second-term recall, there was a main effect of mode of presentation [$F(1,221) = 8.96$]. Single-pair presentation led to greater recall ($M = .067$) than did simultaneous presentation ($M = .049$). More importantly, an effect of difficulty was observed [$F(2,221) = 31.41$]. The probability of recalling second terms alone increased as difficulty increased, with means equal to .023, .068, and .085 for the read, easy completion, and hard completion groups, respectively. The interaction between mode of presentation and encoding difficulty was reliable [$F(2,221) = 9.76$]. The MS_e for

Table 3
Summary of Results from Experiment 3 as a Function of Mode of Presentation and Difficulty

Measure	Simultaneous			Single-Item		
	Read	Easy	Hard	Read	Easy	Hard
Completion		.997	.980		.984	.949
Recall						
Pairs	.504	.498	.439	.439	.353	.346
First	.029	.027	.038	.032	.033	.037
Second	.033	.054	.061	.013	.083	.112
Phi	.654	.625	.577	.629	.495	.464
Categories	.913	.861	.814	.850	.755	.727
Ratio of repetition	.255	.289	.197	.273	.206	.232
<i>n</i>	41	38	36	40	39	33

all tests was .0026. By inspecting Table 3, one can see that the positive generation effect was greater in the single-pair presentation condition than in the simultaneous mode of presentation.

Categories recalled. The probability of correctly recalling categories was reliably affected by mode of presentation [$F(1,221) = 19.03$] and by encoding difficulty [$F(1,221) = 11.28$]. Simultaneous presentation ($M = .86$) led to greater category recall than single-pair presentation ($M = .78$). Category recall declined as encoding difficulty increased ($M_s = .88, .81, \text{ and } .77$ for the read, easy completions, and hard conditions, respectively). The interaction between mode of presentation and difficulty was not reliable ($F < 1.0$). The MS_e for all tests was .0215.

Ratio of repetition. The ratio of repetition declined as encoding difficulty increased, with means equal to .26, .25, and .21 for the read, easy encoding, and hard encoding conditions, respectively. However, this effect of encoding difficulty only approached significance [$F(2,221) = 2.46, p = .088$]. The interaction between mode of presentation and encoding difficulty was reliable [$F(2,221) = 4.41$]. The effect of mode of presentation was minimal [$F(2,221) = .28$]. In order to evaluate further the effects of generation on clustering, comparisons were made between the performance of the easy and hard generation conditions combined and the read condition (Winer, 1971, p. 215). Because there was an interaction between encoding difficulty and mode of presentation, these comparisons were made separately for each mode of presentation. For the simultaneous mode, the comparison was not reliable [$F(1,112) = .19$]. However, for the single-pair mode of presentation, reading led to greater clustering ($M = .273$) than did the two generation groups combined [$M = .219; F(1,109) = 3.94$]. The MS_e for all tests was .0186.

Phi. The measure of dependency between first- and second-term recall indicated a main effect of mode of presentation [$F(1,221) = 25.30$]. Simultaneous presentation led to greater dependency in recall ($M = .62$) than did single-item presentation ($M = .53$). In addition, there was a main effect of encoding difficulty [$F(2,221) = 16.34$], with means equal to .64, .56, and .52 for the read, easy completion, and hard completion groups, respectively. The interaction between mode of presentation and encoding difficulty was also significant [$F(2,221) = 3.45$]. This interaction was the result of a larger effect of encoding difficulty in the single-item mode of presentation than in the simultaneous mode (see Table 3). The MS_e for these tests was .0178.

Discussion

Several important conclusions can be made on the basis of Experiment 3. First, the effects observed in Experiments 1 and 2 cannot be attributed to the mode of presentation. In fact, several of the findings reported above were larger in the single-item mode than in the simultaneous mode. Most notably, the negative effects of generation

on the ratio of repetition and the phi measure of association were larger in the single-item mode of presentation than in the simultaneous mode. Also, the positive effect of generation on second-term recall was larger in the single-item mode than in the simultaneous mode. Independent of mode of presentation, pair recall and category recall declined as encoding difficulty increased.

The second major result of Experiment 3 comprised the effects of generation on organization of pair recall. Category recall and the ratio of repetition declined as encoding difficulty increased. Thus, increases in the encoding difficulty of the second word in a word pair led to decreases in processing across pairs. The ratio-of-repetition data were disappointing in that they did not more strongly support the category-recall results. The sensitivity of the measures of organization may have been impaired due to the fact that some subjects recalled partial pairs mixed in with the intact pairs. This is an unavoidable problem when scoring free recall of pairs.

The third major result was the replication of the negative effects of generation on within-pair processing. As in Experiment 2, the dependence between first- and second-term recall as measured by phi was negatively affected by generation. In addition, these results confirmed the tradeoff between the individual-item processing of the second term and both within- and between-pair relational processing. Increases in the difficulty of the generation task led to an increase in individual-item processing of the second term. This increase in individual-item processing occurred at the expense of relational processing.

The effects of mode of presentation complemented the effects of encoding difficulty in that both single-item presentation and generation led to greater second-term recall than did simultaneous presentation and reading, respectively. Apparently, generation and single-item presentation encouraged processing of the second terms alone. In contrast, pair recall, category recall, and phi were lower in the single-item and generation conditions than in the simultaneous and read conditions, respectively. These results lend further support to the hypothesized tradeoff between the processing of the second term and relational processing both within and between pairs.

EXPERIMENT 4

In each of the above experiments, word presentation was subject-paced. The subjects in the read groups might have spent more time studying each word than the subjects generating the items, leading to better memory in the read groups. In addition, as noted above, several subjects had to be dropped from Experiment 3 because they did not finish the generation task in the allotted time. Thus, the negative generation effects observed above might only have been obtained because of subject-paced presentation. This interpretation of the results seemed unlikely, however, because Slamecka and Graf (1978) had reported identical positive effects of generation with experimenter- and self-paced presentations. In Experi-

ment 4, a single-item, experimenter-paced presentation was used to demonstrate that the negative generation effect occurs independently of mode of presentation.

In Experiment 4, we also addressed a second issue. In each of the experiments reported above, the subjects were instructed to study and recall word pairs. The subjects might have first attempted to recall the first member of each pair, and then used the recalled term as a cue to recall the second member. Generation might have a negative effect on first-term recall (contrary to the results of Experiments 2 and 3), leading to reduced pair recall. This recall strategy provided at best a partial explanation of the results of the first three experiments. However, it was still important to determine whether the negative generation effect was dependent on the instructions to recall pairs. In Experiment 4, two instructional groups were crossed with a read and a generation group. In the *pairs* conditions, the subjects were instructed as in Experiments 1, 2, and 3, and thus were asked to study and recall pairs of words. In the *all items* conditions, the subjects were told to try to learn and recall all the words in their booklets without regard to how they were paired.

Method

Subjects and Design. A total of 63 subjects served in 4 different experimental conditions. Two levels of difficulty of encoding (read and hard completion) were combined with two instructional groups (*pairs* vs. *all items*) in a between-subjects design. The number of subjects in each cell is reported in Table 4.

Materials. The word lists employed in Experiment 3 were again used. Booklets were constructed, with one word pair on each page. The second term in each word pair was either presented intact (read condition) or with several letters missing (generation condition). Words to be generated were presented in the form that they had in the hard generation condition of Experiment 3.

Procedure. With the following exceptions, the procedure closely followed the procedures of Experiments 1, 2, and 3. Prerecorded tones sounded every 5 sec to pace the subjects through their booklets. Subjects in the *pairs* conditions were told to try to recall the word pairs, and if they could not remember one of the words from a word pair, to record the individual word they did remember. These instructions were the same as the instructions used in Experiments 2 and 3. In the *all items* conditions, the subjects were asked to recall all the words in the booklet freely, and they were told, "It does not matter if you can only remember one word from a pair, or the position of the words in pairs, rather, just try to recall all the words." These recall instructions were given prior to presentation as well as prior to the recall test.

Results

The instructional manipulation had a major impact on the recall protocols. To assess this impact, we calculated the number of subjects who recorded both members of every recalled pair together on the same line when they recalled word pairs. Under the *pairs* instructions, 26 subjects (81%) recorded pairs together. With the *all items* instructions, only 6 subjects (19%) recorded the pairs together. These differences provided strong evidence that the instructional manipulation was successful. However, given these vast differences in recall protocols, comparable measures of category organization in recall could not

be calculated for both instructional groups. For this reason, the number of categories recalled and the ratio-of-repetition measure reported in Experiment 3 were not analyzed in Experiment 4. We did calculate the proportion of intact pairs recalled, first terms recalled only, second terms recalled only, and the phi measure of dependence in pair recall. Each of these measures is discussed below. A summary of the results is presented in Table 4.

Completion. The mean completion rate was .97. The recall instructions did not reliably influence completion [$F(1,29) = .67$, $MS_e = .0008$].

Pairs recalled. Given the above-noted differences in the recall protocols, the subjects were given credit for recalling a pair as long as both members of the pair appeared on the recall sheet. Surprisingly, there was no effect of instructions on the proportion of pairs recalled [$F(1,59) = .78$]. There was an effect of encoding task [$F(1,59) = 33.30$]. The read task led to greater recall ($M = .40$) than did the generation task ($M = .31$). The instruction \times task interaction was not significant [$F(1,59) = .70$]. The MS_e for these tests was .0117.

First-term recall. There was no effect of instructions [$F(1,59) = 2.09$] or task [$F(1,59) = .13$] on the proportion of first terms recalled alone. The instruction \times task interaction was also nonsignificant [$F(1,59) = .43$]. The MS_e for these tests was .0014.

Second-term recall. A greater proportion of second terms was recalled alone with the *all items* instructions ($M = .11$) than with the *pair* instructions [$M = .07$; $F(1,59) = 4.69$]. In addition, second-term recall was greater with the generation task ($M = .15$) than with the read task [$M = .03$; $F(1,59) = 49.10$]. The interaction between instruction and task was not significant [$F(1,59) = .34$]. The MS_e for these tests was .0045.

Phi. The only significant effect on phi was the negative generation effect [$F(1,59) = 52.10$], in which reading led to greater dependence between first- and second-term recall ($M = .58$) than generation did ($M = .35$). The effects of instructions [$F(1,59) = 2.19$] and the task \times instructions interaction [$F(1,59) = .16$] were not reliable. The MS_e for these tests was .0168.

Discussion

The results of Experiment 4 once again demonstrated the negative effects of generation on pair recall and on pair integration. The major purpose of this experiment was to rule out several uninteresting interpretations of the

Table 4
Summary of Results from Experiment 4 as a
Function of Task and Recall Instructions

	Pairs		All Items	
	Read	Generate	Read	Generate
Pairs recalled	.420	.239	.373	.238
First terms	.025	.034	.045	.042
Second terms	.020	.129	.047	.175
Phi	.612	.363	.551	.328
<i>n</i>	16	16	16	15

negative generation effect. One interpretation was that the negative generation effect was the result of special study strategies adopted in a self-paced mode of presentation. The negative generation effects obtained with experimenter-paced presentation ruled out this interpretation. If one is willing to hazard between-experiment comparisons, one can compare the magnitude of the positive and negative generation effects (read vs. hard generation) in the single-item mode of presentation in Experiment 3 to the magnitude of these effects in Experiment 4. The negative generation effect on pair recall appeared more robust in Experiment 4 ($M = .16$) than in Experiment 3 ($M = .09$). Similarly, the negative generation effect on phi was greater in Experiment 4 ($M = .24$) than in Experiment 3 ($M = .17$). In contrast, the positive effects of generation on second-term recall were comparable in Experiments 4 ($M = .12$) and 3 ($M = .10$).

A second uninteresting interpretation of the negative generation effect is that it was the result of the subjects' trying to recall intact word pairs. Instructions directing the subjects not to worry about pairing the items in recall successfully produced disorganized recall protocols. The instructions to recall all items also increased the number of second terms recalled alone. However, the instructions had no impact on the magnitude of either the positive or negative generation effects observed.

EXPERIMENT 5

In Experiments 1-4, positive effects of generation were found on the recall of the second terms. Our interpretation of this result was that generation led to greater individual-item processing of the second member of the word pairs than did reading. A direct implication of this hypothesis is that recognition of second terms should be greater following generation than following reading. This prediction has received support in previous studies (e.g., Slamecka & Graf, 1978). The purpose of Experiment 5 was to replicate the positive effects of generation on second-term recognition with the materials and procedures used in the first four experiments.

Method

Subjects and Design. Seventy-eight subjects from the same population as was drawn from in Experiments 1-4 served in this experiment. Two conditions were tested in a between-subjects design: a read condition ($N = 40$) and a generate condition ($N = 38$).

Materials. The materials used in Experiment 2 were again employed. The word frequencies of the second terms were determined, using the Thorndike and Lorge (1944) word count. For each second term, three unrelated words matched in frequency were selected from Thorndike and Lorge. These words served as distractors on a four-alternative forced-choice recognition test. The first terms from the word pairs were not provided on the test form. The word pairs were randomly ordered for presentation and the sets of test items were randomly ordered on a single page for the recognition test.

Procedure. The procedure was nearly identical to the procedure for the associate, simultaneous, self-paced groups in the above experiments. That is, the subjects were asked to associate the words in each pair. They were told that they would be asked to recall the

word pairs, equating test expectancy across experiments. The procedure for Experiment 5 was different from that of Experiments 1-4, in that just prior to the memory test, the subjects were informed about the nature of the recognition test. The instructions stressed that the subjects were to circle one item on each line, even if they did not know the correct answer.

Results and Discussion

The probability of correct completion for the generation group was .97 items. The generation task led to a higher probability of correct recognition ($M = .988$) than did reading [$M = .934$; $F(1,76) = 15.92$]. This result replicates previous research, demonstrating a positive effect of generation on recognition in a between-list design with our materials and procedure.

GENERAL DISCUSSION

The mnemonic consequences of reading, performing an easy generation task, and performing a hard generation task were evaluated in between-list designs in five experiments. In between-list designs, the effects of generation cannot be attributed to differences in attention or priority given to the generated items at the expense of read items. Reading the material led to greater free recall of word pairs, greater cued recall of the second word from weak associates, greater pair integration, and greater organization in recall, when compared with generating the second word in the pair. As the generation task became more difficult, pair recall, cued recall, pair integration, and organization in recall declined. The negative generation effect was found with strong, moderate, and weak associates; with a categorized word list; with instructions to copy the items, to associate the items, and to recall items without regard to pairings; and with three modes of presentation (self-paced simultaneous, self-paced single-item, and experimenter-paced single-item). Generating the second member of a word pair had a positive effect on the recognition of that member, and positive effects on the probability of recalling the second member alone.

One interpretation of the results reported above is that, in between-list designs, generation inhibits both within- and between-pair relational processing. In contrast, generation leads to greater individual-item processing of the second word when compared with reading. Increases in the difficulty of the generation task lead to increases in individual-item processing at the expense of relational processing. This interpretation is more than just a restatement of the results, in that relational and individual-item processing were operationalized in a number of different ways. Relational and individual-item processing were manipulated with instructions (associate vs. copy) and with mode of presentation. The ratio-of-repetition measure of organization and phi provided different measures of relational processing. Finally, second-term recall and recognition provided measures of individual-item processing.

A potential criticism of the research reported above is that it has little to do with the typical generation effect

since pure comparisons between read and generate conditions were not made. In each experiment, the subjects were asked to associate the words in the word pairs. In response to this criticism, we would like to mention three things. First, in Experiment 2 we specifically manipulated instructions. The negative effects of generation we observed were found with both associate and copy instructions. Second, as noted above, instructions similar to our "associate" instructions were employed by Rabinowitz and Craik (1986, Experiment 3). Third, with few exceptions, the subjects in generation experiments have always been asked to perform some task in addition to reading and generating the items. Gardiner et al. (1985), Graf (1982), Jacoby (1978), Payne, Neely, and Burns (1986), Slamecka and Fevreski (1983), and Slamecka and Graf (1978) asked subjects to read aloud the second word. Donaldson and Bass (1980), Nairne, Pusey, and Widner (1985), and Nairne and Widner (1987) asked subjects in the read condition to copy the second word. In contrast, Glisky and Rabinowitz (1985) and McDaniel (1984) had subjects read the material to themselves. Nonetheless, in each of these experiments, positive generation effects were reported. Thus, it was not our instructions that led to the negative generation effect, but the experimental design and the type of memory test employed.

Why has no one else reported negative effects of generation? There are a number of possibilities. First, most researchers have employed within-list designs in which generated items may receive attention at the expense of read items. Thus the research presented above adds further weight to Poulton's (1982) criticism that cognitive psychologists rely too heavily on these within-subjects designs. Second, the negative effects reported above were rather small, ranging from 2% to 18%. In comparison, the positive effects of generation obtained in within-list designs are relatively large and easily detected. When researchers have employed between-list designs, *ns* appropriate for within-list comparisons have been used, and consequently no effects of generation have been detected. In contrast, because we had difficulty finding a positive generation effect, in each of our experiments we employed relatively large *ns* with as many as 40 subjects per cell. Third, the negative effects of generation on relational processing occurred conjointly with positive effects on individual-item processing. The net generation effect may depend on the extent to which relational and individual-item processing contribute to performance on a given type of memory test (McDaniel et al., 1986). The free recall of word pairs we employed as a memory test may be particularly sensitive to relational processing, and thus sensitive to the negative effects of generation. Fourth, two very recent studies have reported negative generation effects. In between-list comparisons, Begg and Snider (1987) found negative generation effects in cued recall with unrelated pairs. However, the subjects were first required to make a recognition judgment of the cue prior to attempting to recall the second item in the pair. With a more traditional recall test, Hirshman and Bjork (1988,

Experiments 1 and 2) also reported negative generation effects. However, their effects were not significant (they employed 20 subjects per cell).

Hirshman and Bjork (1988) evaluated the effects of generation on individual-item and relational processing. As in the experiments reported above, Hirshman and Bjork employed between-list designs. Unlike the results of our experiments, larger positive effects of generation were found in cued recall than free recall. A number of factors distinguish our experiments from those reported by Hirshman and Bjork. First, Hirshman and Bjork focused on within-pair processing, and thus their experiments may not have been sensitive to the negative effects of generation on between-pair relational processing reported above. Second, in their free recall task, the subjects were asked to recall only the second terms from the word pairs. As noted above, we also found positive effects of generation on second-term recall. In addition, as noted above, Hirshman and Bjork did find some evidence for negative effects of generation on free recall. Considering these factors, we do not find our results inconsistent with those reported by Hirshman and Bjork.

The small negative generation effects we reported may strike some as uninteresting. However, the direction of the effects is of more interest than their magnitude. Given that generation may actually hurt memory performance, great caution must be observed before the generation effect is employed in real-world settings. Also, the negative generation effect may challenge the view that increases in encoding difficulty are associated with improved memory performance (Tyler et al., 1979). Thus the negative generation effect is of both practical and theoretical importance.

REFERENCES

- BATTIG, W. F., & MONTAGUE, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, **80**(3, Pt. 2).
- BEGG, I., & SNIDER, A. (1987). The generation effect: Evidence for generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 553-563.
- BOUSFIELD, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, **49**, 229-240.
- CRAIK, F. I. M., & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, **11**, 671-684.
- DONALDSON, W., & BASS, M. (1980). Relational information and memory for problem solutions. *Journal of Verbal Learning & Verbal Behavior*, **19**, 26-35.
- EINSTEIN, G. O., & HUNT, R. R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 588-598.
- GARDINER, J. M., SMITH, H. E. C., RICHARDSON, C. J., BURROWS, M. V., & WILLIAMS, S. D. (1985). The generation effect: Continuity between generating and reading. *American Journal of Psychology*, **98**, 373-378.
- GLISKY, E. L., & RABINOWITZ, J. C. (1985). Enhancing the generation effect through repetition of operations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 193-205.

- GRAF, P. (1982). The memorial consequences of generation and transformation. *Journal of Verbal Learning & Verbal Behavior*, **21**, 539-548.
- 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.
- HUNT, R. R., & EINSTEIN, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior*, **20**, 497-514.
- HUNT, R. R., & SETA, C. E. (1984). Category size effects in recall: The roles of relational and individual-item information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 454-464.
- 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.
- MCDANIEL, M. A. (1984). The role of elaborative and schema processes in story memory. *Memory & Cognition*, **12**, 46-51.
- MCDANIEL, M. A., EINSTEIN, G. O., DUNAY, P. K., & COBB, R. C. (1986). Encoding difficulty and memory: Toward a unifying theory. *Journal of Memory & Language*, **25**, 645-656.
- NAIRNE, J. S., PUSEN, C., & WIDNER, R. L., JR. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition*, **13**, 183-191.
- NAIRNE, J. S., & WIDNER, R. L., JR. (1987). Generation effects with nonwords: The role of test appropriateness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 164-171.
- PALERMO, D. S., & JENKINS, J. J. (1964). *Word association norms*. Minneapolis: University of Minnesota Press.
- PAYNE, D. G., NEELY, J. H., & BURNS, D. J. (1986). The generation effect: Further tests of the lexical activation hypothesis. *Memory & Cognition*, **14**, 246-252.
- POULTON, E. C. (1982). Influential companions: Effects of one strategy on another in the within-subject designs of cognitive psychology. *Psychological Bulletin*, **91**, 673-690.
- RABINOWITZ, J. C., & CRAIK, F. I. M. (1986). Specific enhancement effects associated with word generation. *Journal of Memory & Language*, **25**, 226-237.
- SLAMECKA, N. J., & FEVREISKI, J. (1983). The generation effect when generation fails. *Journal of Verbal Learning & Verbal Behavior*, **22**, 153-163.
- 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.
- THORNDIKE, E. L., & LORGE, I. (1944). *The teachers' word book of 30,000 words*. New York: Teachers College Press, Columbia University.
- TYLER, S. W., HERTEL, P. T., MCCALLUM, M. C., & ELLIS, H. C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 607-617.
- WINER, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

(Manuscript received January 26, 1988;
revision accepted for publication September 6, 1988.)