Journal of Experimental Psychology:
Learning, Memory, and Cognition
1991, Vol. 17, No. 4, 702-709

Copyright 1991 by the American Psychological Association, Inc.
0278-7393/91/$3.00

Dissociative Effects of Generation on Item and Order Retention

James S. Nairne, Gregory L. Riegler, and Matt Serra
Purdue University

The effects of generation on the long-term retention of item and order information were examined in a between-list design in 3 experiments. In each experiment, completing word fragments during presentation significantly impaired long-term retention of serial order, as measured by either a reconstruction task or the amount of input-output correspondence in free recall. Memory for the individual items, however, was sometimes helped by generation. This pattern of dissociation, reminiscent of immediate memory findings, is used to interpret problematic issues in the generation effect literature and to argue for the role of the item-order distinction in the long-term-memory arena.

Memory researchers generally agree that the process of single-trial recall, either free, cued, or serial, depends on many kinds of mnemonic information. Knowledge about how list items are related, for example, can be used to generate output candidates at the time of test; individual-item information, uniquely defined by the encoding context, can make it easier to pick target items in the candidate set from distractors that did not occur in the study episode (e.g., Bellezza, Cheesman, & Reddy, 1977; Hunt & Einstein, 1981; Mandler, 1980). From a diagnostic standpoint, however, interpretative problems can arise because a single independent variable can affect these various kinds of mnemonic information in quite different ways.

Examples of such dissociative effects abound in the memory literature but have been displayed clearly and with theoretical import in studies of immediate retention. When items are drawn from the same taxonomic category, for instance, subjects show an enhanced ability to remember the specific items that occurred on a trial, but an impaired ability to remember the order in which those items were presented. Dissociative effects of this type, among others (see Bjork & Healy, 1974; Healy, 1974, 1982; Murdock, 1976; Murdock & Vom Saal, 1967), convinced researchers studying immediate memory to use separate indexes of item and order memory and not to rely on a single measure of performance as the cornerstone of theory (Drewnowski, 1980; Estes, 1973; Murdock, 1983; Nairne, 1990a; Shiffrin & Cook, 1978). In the present context, we apply similar logic to a long-term-memory phenomenon, the generation effect (Slamecka & Graf, 1978), in which internally generated material, typically, is better remembered than is information supplied by the experimenter. On the basis of a sampling of recent studies (e.g., Begg & Snider, 1987; Schmidt & Cherry, 1989; Slamecka & Katsaiti, 1987), we suspected that stimulus generation might produce dissociative effects on item and order memory and that these dissociations might account for some of the recent controversies in the generation effect literature.

The idea that generation might produce opposite effects on item and order memory has been suggested in prior studies but has never been tested directly. Consider the finding that the generation effect is reduced or eliminated in the free recall of lists containing unrelated words when reading and generating occur between subjects or in different lists (Begg & Snider, 1987; Hirshman & Bjork, 1988; McDaniel, Waddill, & Einstein, 1988; Slamecka & Katsaiti, 1987). Under such conditions, generation might conceivably lead to a variety of effects, including enhanced item memory but impaired memory for order. Because subjects often use seriation, recall based on serial order as an output strategy (e.g., see Kintsch, 1970; Mandler, 1969; Postman, 1972), an impairment in memory for the relative orderings of the items might well mask any increase in the ability to remember the items themselves, reducing the apparent size of the generation effect. In contrast, when a categorized list is used, in which a form of relational information is built into the generic structure of the list, retention of serial order information might not be critical because alternative output strategies are available (e.g., category generation). In this case, one might anticipate an unmasking of the improved item memory and a generation effect in recall; such a result has recently been obtained by McDaniel et al. (1988) but interpreted from a quite different perspective.

The picture has been greatly complicated by the failure of researchers to use retention tasks that directly measure these different kinds of mnemonic information. Typically, one finds that only recall or recognition are used, and attempts to isolate organizational effects are derived from the presence of clustering patterns in recall protocols. Although interesting in principle, these clustering effects can be small and variable (see McDaniel et al., 1988); moreover, clustering analysis is appropriate only when categorized lists are used, which limits its applicability as a theoretical tool. In the present set of experiments, we took a more direct approach by providing a

We would like to thank Ian Begg, John Gardiner, and Elliot Hirshman for their constructive comments on the manuscript. Thanks are also due to Stephany Bonine, Melynda Eastman, Tonia Osberg, John Salapski, Tate Stuntz, Christy Wheeler, and Karen Whelan for help in running subjects.

Gregory L. Riegler is now at the University of St. Thomas, St. Paul, Minnesota.

Correspondence concerning this article should be addressed to James S. Nairne, Department of Psychology, Purdue University, 1364 Psychology Building, Room 3152, West Lafayette, Indiana 47907.

measure of order memory that was independent of recall. We reasoned that any decrement in order memory, as a function of generation, would be a key piece in the puzzle surrounding the presence or absence of the generation effect in free recall.

To provide a pure measure of serial order information, it is necessary to control for subjects' abilities to remember the items themselves. Such control can be achieved in several ways, but one preferred method is to use a reconstruction task (Crowder, 1979; Healy, 1974; Horowitz, 1961; Nairne, 1990b). At test, subjects are given the items from the list and are asked to place them in their original order of presentation. Because subjects are fully informed that the items are exclusively from the list, it is usually assumed that a pure form of order, or possibly position, memory is being tested. Such a measure contrasts with straightforward recall (either serial or free) in which memory for the items and for the order in which they occurred is confounded. Because it is also of interest to replicate the effect of generation on the more typical measures, subjects either recognized or recalled the items across the three experiments. As we shall show, generation had quite different effects on the various retention measures.

Experiment 1

Our procedures differed somewhat from those normally used in generation studies. Rather than using a single long list, subjects received a number of short lists that were tested after a 30-s period of distraction or in a final end-of-session test. On a random half of the trials, each of the eight items in a list was presented in a fragment form, with one letter missing, requiring generation; on the other half of the trials, the items were presented intact. Each list was followed by the distraction period but ended unpredictably with either a reconstruction task requiring the reproduction of serial order or simply a line of asterisks. At the end of the session, all of the subjects received a surprise yes/no recognition test for the read and generated items from the untested lists.

Note that this design uses a between-list manipulation of reading versus generating. Consequently, any obtained generation effect, in either item (final recognition) or order (reconstruction) retention, is not easily interpreted from a selective rehearsal perspective (see Slamecka & Katsaiti, 1987). It seems highly unlikely that subjects would choose to rehearse generated items selectively because both read and generated items are presented and tested independently on completely different trials. With respect to other accounts of the generation effect (see Begg, Snider, Foley, & Goddard, 1989; Slamecka & Katsaiti, 1987, for reviews), predictions at this point would be post hoc because none of these accounts has explicitly considered a role for serial order information.

Method

Subjects and apparatus. Subjects were 32 undergraduates, who participated for course credit in introductory psychology. Individual sessions were conducted, each lasting approximately 1 hr. All of the stimulus materials were presented and controlled by an IBM-compatible computer.

Materials. The 192 stimulus items were medium- and high-frequency nouns, four to six letters in length, drawn from the Paivio, Yuille, and Madigan (1968) norms. When presented in fragment form, one letter was deleted from the stimulus and replaced with an underline character. The particular letter was selected according to two criteria: (a) The solution had to be relatively obvious to reduce generation failures, and (b) the fragment could have only one solution. Each list contained 8 unique words, which were completely counterbalanced across the major conditions of interest. For example, each list was presented in the read and generate form, across subjects, and was tested in either reconstruction or final recognition. When an individual list was created, the stimulus items were randomly assigned to each of the eight serial positions.

Procedure. Each session presented 24 eight-item lists, which were preceded by 4 practice lists and followed by a single recognition test. Subjects were fully informed about the nature of the reconstruction task but were given no information about the final recognition test.

Trials began with the word *ready*, displayed for 1,000 ms and accompanied by a short beep. The eight stimuli were then presented, in either read or generate form, centered in the middle of the terminal screen. Each item appeared for 2,000 ms, with a 500-ms interstimulus interval. Subjects were asked to repeat an item aloud when it appeared, generating if necessary, and any generation failures were noted by the experimenter. After the last item on the list, single digits, drawn randomly from the set 0–9, began to appear in succession. Subjects were required to press one key on the numeric keypad if the digit was even and another, adjacent key if the digit was odd. Subjects read the digits silently; each digit remained on the screen for 1,000 ms. This distractor task continued for 30 s and was followed, on a random half of the read and generate trials, by either (a) the reconstruction test or (b) a line of asterisks. Subjects had no way of predicting what kind of test would follow the distraction period.

For the reconstruction test, all of the eight items from the trial were presented simultaneously, in a line covering the middle of the terminal screen, but in a new random order. The subject's task was to write the items on a sheet of paper in their original order of presentation. The paper contained eight spaces for the responses next to each of the numbers 1–24. Subjects were instructed to fill in each of the response blanks and not to repeat an item. The test was subject-paced; the items remained on the screen until the space bar was pressed to initiate the next trial. When the line of asterisks appeared, rather than the reconstruction test, subjects were told simply to wait for the next list that began immediately after the asterisks disappeared from the screen. The line of asterisks remained on the screen for 5 s.

After the 12 read and 12 generate trials were concluded, the final recognition test was administered. Subjects were asked to make yes/no decisions about prior occurrence for 144 items (96 targets and 48 distractors), presented individually on the CRT (cathode-ray tube) screen. The target items consisted of all of the read and generated items from the untested lists (8 items from 6 read lists and 8 items from 6 generated lists) and were randomly intermixed with distractors that were matched on stimulus characteristics. Each item remained on the screen until the subject responded by pressing one of two keys. Immediately after a decision was made, the next item was presented for the recognition judgment. Three 5-s breaks were given, signaled by a line of asterisks, at equal intervals throughout the test.

Results and Discussion

Subjects correctly generated the presented item on 93.5% of the generation opportunities, and there were no read errors. The overall patterns did not depend on the success or failure of generation, but at least for the reconstruction data, both the unconditionalized and conditionalized results are reported. Reliability in the statistical sense was measured at $p < .05$.

Reconstruction. The results from the reconstruction task, scored without regard to generation success or failure, are displayed in Figure 1. The results are plotted as mean proportion correct, which provides an indication of how well subjects were able to reposition the items from each of the eight serial positions during the reconstruction test. The overall analysis of variance (ANOVA) on the data, which was conducted on the number correct for each position, revealed reliable effects of serial position, $F(7, 217) = 31.39$ ($MS_e = 1.23$); encoding condition (read vs. generate), $F(1, 31) = 6.02$ ($MS_e = 5.32$); and the Encoding × Position interaction, $F(7, 217) = 2.10$ ($MS_e = 0.88$). The serial position curves showed marked primacy effects and a slightly bow-shaped form; differences between the read and generated lists were absent for the primacy portion of the list, accounting for the interaction. Overall, performance on the reconstruction task was clearly impaired by generation, reflecting poorer memory for serial order. For the 32 subjects, 23 showed better reordering performance on read trials, compared with generate trials; 9 subjects showed the reverse pattern.

Similar conclusions are reached from examination of the conditionalized data, in which generation success during encoding was taken into account. If an item was not generated correctly, it was simply removed from the analysis, and the appropriate proportions were calculated. The ANOVA again revealed reliable effects of serial position, $F(7, 217) = 28.03$ ($MS_e = .04$); encoding condition, $F(1, 31) = 5.30$ ($MS_e = 0.15$); and the Encoding × Position interaction, $F(7, 217) = 2.40$ ($MS_e = 0.03$). Collapsed across serial position, reconstruction performance averaged 0.57 on the read trials and 0.50 on the generate trials. Across the 32 subjects, 21 showed an overall read advantage, 10 showed a generate advantage, and there was one tie. One can rule out, therefore, the idea that the pattern shown in Figure 1 resulted from the failure to generate some of the presented items correctly. Rather, the data indicate that fragment completion, at least under the presentation conditions adopted here, can disrupt a subject's ability to reconstruct the original order of presentation in a list.

Recognition. The final recognition test was designed to get a measure of individual item memory under conditions in which no memory for serial order information was required. On the basis of past studies of the generation effect, it was expected that generation would improve long-term recognition of the item's occurrence in the list (see Slamecka & Graf, 1978), although in a between-list design the size of the effect may be reduced (see Begg & Snider, 1987). The unconditionalized data are presented in Figure 2 as a function of encoding condition and serial position. An ANOVA on the number correct revealed a highly reliable effect of encoding condition, $F(1, 31) = 23.72$ ($MS_e = 1.59$), as well as a significant effect of serial position, $F(7, 217) = 3.42$ ($MS_e = 1.29$); the interaction did not approach significance ($F < 1$), and the false alarm rate was quite low (.061). The presence of a reliable generation effect in the recognition data, although not surprising, is important because it indicates that the order decrement cannot be attributed simply to poor individual-item memory (e.g., a failure to encode generated items adequately).

The results of Experiment 1 fit a clear pattern of dissociation, in which a single independent variable shows an opposite, crossover pattern on two different dependent variables (see Roediger, 1984). The fact that generation produces such results, which depend on the retention measure adopted, was not unexpected (see Jacoby, 1983; Nairne, 1988; Schmidt & Cherry, 1989); in this case, however, the dissociation of item and order retention is relevant to recent controversies about the basis of the generation effect. If generation impairs a subject's ability to remember the serial order of list presentation, as the present results document, then it is natural to expect that free recall would also be adversely affected. If subjects rely on seriation as an output strategy, then generated items might receive less consideration as output candidates, even though it would be relatively easy for the subjects to identify them as targets. To provide further evidence on this point, Experiment 2 was designed to provide a measure of seriation in free recall, as a function of generation, and to replicate the ordering effects of Experiment 1. On the basis of the present findings, we expected that subjects would show

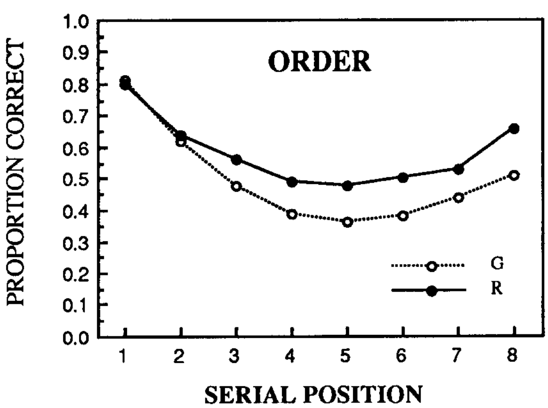

Figure 1. Reconstruction performance for the read (R) and generated (G) lists in Experiment 1, plotted as a function of serial position.

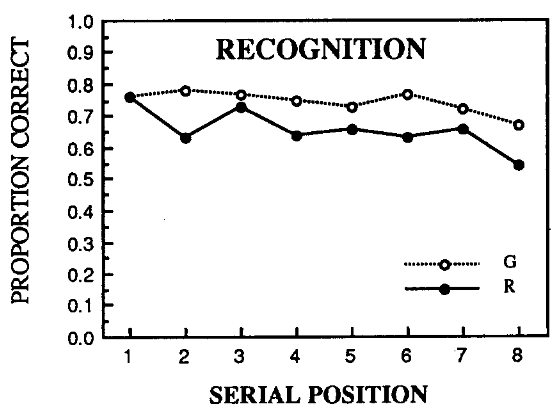

Figure 2. Recognition performance for the read (R) and generated (G) lists in Experiment 1, plotted as a function of serial position.

less input–output correspondence in recall when lists were generated, as opposed to read.

Experiment 2

Experiment 2 was a direct replication of Experiment 1, except that one half of the time trials ended with instructions to free recall the items without regard to serial order. As before, eight-item lists were presented, in either a read or generate form, followed by a 30-s distractor task. The distraction period ended with either a representation of the eight items for reconstruction or the instructions for free recall. Subjects had no way of predicting, on a given trial, whether reordering or recall would be required.

Method

Subjects and apparatus. Subjects were 24 undergraduates, who participated for course credit. Individual sessions were controlled by an IBM-compatible computer.

Materials. The stimulus materials used in Experiment 1 were used again in Experiment 2.

Procedure. The same procedure was used, except that on a random half of the trials, equally divided between read and generate lists, the message "recall the words in any order" appeared in the center of the screen. Subjects were told at the beginning of the experiment that when this message appeared, they were to write down the items from the list in any order. Subjects wrote their responses in each of eight blanks on the answer sheet, as on the reconstruction trials. Once again, subjects had no way of knowing until after the completion of the distraction period whether the recall message or the eight list items for reconstruction would appear. Both retention tests were subject-paced.

Results and Discussion

Correct generation was again quite likely, occurring on 94.2% of the generation opportunities, and there were no read errors.

Reconstruction. The results from the order test are shown in Figure 3. These are the unconditionalized data; the pattern resembles the one shown in Figure 1. The ANOVA revealed significant effects of serial position, $F(7, 161) = 41.25$ ($MS_e = 1.12$), and encoding condition, $F(1, 23) = 11.84$ ($MS_e = 3.60$), but no reliable interaction in this case, $F(7, 161) = 1.88$ ($MS_e = 1.79$). Across the 24 subjects, for the unconditionalized data, 19 showed poorer reconstruction performance on generate trials and 5 showed the reverse pattern. The conditionalized data showed an identical pattern and are not reported here. As in Experiment 1, it is quite clear that generation failed to promote any retention advantage when the test was one of order, rather than item, memory.

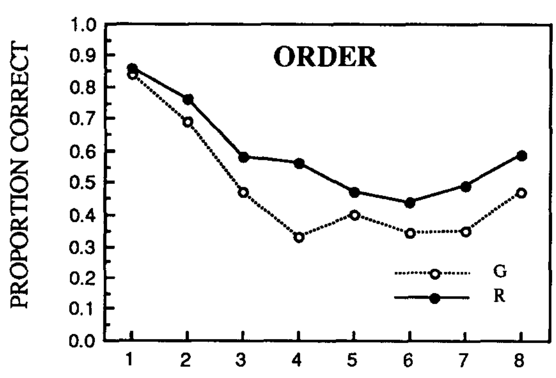

Figure 3. Reconstruction performance for the read (R) and generated (G) lists in Experiment 2, plotted as a function of serial position.

Free recall. The results from the free-recall trials, displayed in Figure 4, also show a clear pattern. These data were conditionalized on successful generation, which seemed more appropriate given that the items were not presented at test (as in reconstruction). The ANOVA revealed highly reliable main effects of serial position, $F(7, 161) = 9.49$ ($MS_e = .046$), and encoding condition, $F(1, 15) = 10.03$ ($MS_e = .03$); the Encoding × Position interaction was not reliable, $F(7, 105) < 1$. Generation had a substantial negative effect on free recall, which is opposite to the standard generation effect but not unprecedented in the literature. Schmidt and Cherry (1989) found, for example, that generating the response member of a word pair led to poorer recall of the pair, given that reading and generating were done in separate lists. They interpreted their "negative generation effect" as resulting from poor stimulus–response integration, which is consistent with the kind of reasoning adopted here. If subjects relied on serial order as an output strategy, which seems reasonable given that serial order was tested on one half of the trials, then impaired serial order memory should have had a negative impact on recall.

To provide a more direct test of this hypothesis, we measured the amount of seriation, or input–output correspondence, across the read and generate trials. We used a technique developed by Asch and Ebenholtz (1962) that provides a relative index of correct input–output correspondence. To use their example, consider that a subject recalled, in order, Items 1, 2, 8, 6, and 3 from the list. If adjacent recalls are considered as pairs, then on this trial the subject recalled four pairs (1–2, 2–8, 8–6, 6–3). The question of interest asks how many of these pairs, relative to the total number recalled, preserve the relative order of input? In this case, two of the four pairs show the correct ordinal sequence, creating an overall index of 0.50. This measure is preferred because it takes into account correct adjacent and remote pairs, as well as the total number of pairs recalled.

In the present case, the average index for read trials was 0.68, compared with 0.62 for generate trials. Both of these indexes are above the chance level of 50% (one half of the pairs in the correct order and one half in the incorrect order), so it seems reasonable to conclude that subjects relied on seriation as an output strategy (see also Bousfield & Abramczyk, 1966; Jahnke, 1965). As with the reconstruction data, the negative effect of generation was statistically reliable, $t(23) = 2.51$. Subjects showed less input–output correspondence when generation was required during initial list presentation. These results are consistent with the general conclusion of Experiments 1 and 2, namely, that generation impairs memory for serial order. Although this analysis does not conclusively demonstrate that the negative generation effect in free recall is due to an ineffective seriation strategy, the Asch–Ebenholtz analysis certainly fits that interpretation.

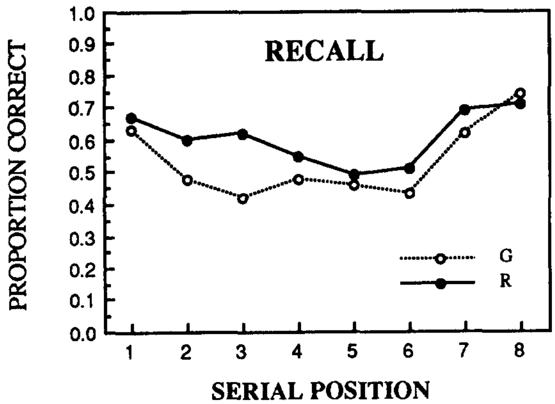

Figure 4. Free-recall performance for the read (R) and generated (G) lists in Experiment 2, plotted as a function of serial position of input.

At this point, a skeptic might argue that the results of Experiment 2 could be more parsimoniously explained by appealing to some sort of general attentional deficit induced by generation. One might expect a very similar pattern of results, reduced recall and reconstruction, for any manipulation that requires some effort during encoding (e.g., a faster presentation rate). Although we concur with elements of this reasoning, the present data require a more complex appeal. The dissociation of Experiment 1 shows that generation produces complex effects, some negative and some positive. It is the relative trade-off between benefit and cost that best characterizes generation and, we feel, demands consideration in any retention test interpretation (see Begg et al., 1989, for a similar point). For example, even though free recall was significantly impaired by generation in Experiment 2, we assume that item-specific processing was enhanced. The positive effects of generation are merely masked because free recall requires organizational processing (e.g., interitem associative links) that is hurt by the fragment completion task. We anticipate, therefore, that if an alternative organizational structure is made available to the subject, positive effects of generation in recall will emerge, even if reading and generating are conducted in different lists. This prediction formed the basis for Experiment 3.

Experiment 3

The design of Experiment 3 resembled Experiment 2 in all respects except that categorized lists were used. Each list contained eight items drawn from the same, but unique, taxonomic category (e.g., flowers). Under these conditions, we would expect the retention of serial order to be harmed by generation, as in the previous two experiments, but we would expect a different pattern in free call. With categorized lists, the inherent category structure provides an alternative organizational scheme (other than seriation): Subjects can simply generate category instances as output candidates and then check their legitimacy as episodic targets. With the unrelated lists of the previous two experiments, organizational structure could only be established at the point of encoding (e.g., linking the items together in a serial chain). We predicted a reversal of the recall pattern of Experiment 2 and a positive generation effect to emerge. We also expected less overall use of seriation as an output strategy in recall for both read and generated lists, although significantly more input–output correspondence should still be present on read trials.

Method

Subjects and apparatus. Subjects were 36 undergraduates, who participated for course credit. All of the subjects were tested in individual sessions on an IBM-compatible computer.

Materials and design. Eight items from 24 different taxonomic categories were drawn from the Battig and Montague (1969) category norms. The first few rated instances from each category were not included to prevent a strategy of simply listing highly probable category members during recall. As in the previous experiments, one letter was deleted from each item to form a readily completed fragment. Items and categories were presented, across subjects, in both read and generated form.

Procedure. The procedure was exactly the same as the one used in Experiment 2. No special instructions were given regarding the use of categories, but because there were 4 practice trials, it is reasonable to assume that subjects were aware that the lists were categorized prior to the start of the 24 critical trials. For the 24 trials, 12 were presented in a read form and 12 were presented in fragment form; 6 of the trials for each encoding condition were followed by reconstruction, and 6 were followed by free recall.

Results and Discussion

Very few generation errors occurred. Subjects correctly completed the word fragments on 98% of the generation opportunities, and there were no read errors.

Reconstruction. The results from the reconstruction trials are plotted in Figure 5, broken down by encoding condition and serial position. In replication of the previous two experiments, the unconditionalized ANOVA revealed significant effects of serial position, $F(7, 245) = 53.29$ ($MS_e = 1.61$), and encoding condition, $F(1, 35) = 8.79$ ($MS_e = 2.26$); the Encoding × Position interaction did not reach conventional levels

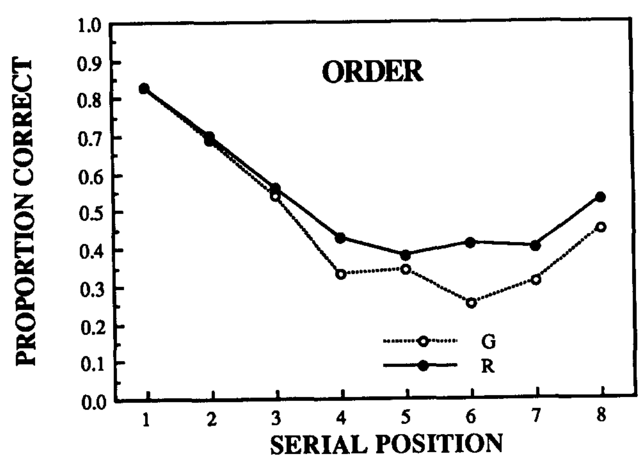

Figure 5. Reconstruction performance for the read (R) and generated (G) lists in Experiment 3.

of significance, $F(7, 245) = 1.82$ ($MS_e = 0.93$). Across subjects, 24 showed better reconstruction performance on read trials, 9 showed a generation advantage, and there were 3 ties. Clearly, despite the use of categorized lists, fragment completion during presentation still interfered with the retention of serial order. This was the expected result because there is nothing inherent in category structure to provide information about the experimenter-determined (random) order of list items. The conditionalized data showed a virtually identical pattern and are not presented here.

Free recall. The results from the free-recall trials, for read and generated lists, are shown in Figure 6 as a function of serial position. For these data, we anticipated a generation effect, but as Figure 6 indicates, an Encoding × Position interaction was the dominant feature of the pattern. An overall ANOVA on the recall data indicated significance only for position, $F(7, 245) = 9.75$ ($MS_e = 0.03$), and the Encoding × Position interaction, $F(7, 245) = 2.30$ ($MS_e = 0.03$); the main effect of encoding condition was not reliable, $F(1, 35) = 2.52$ ($MS_e = 0.03$). To analyze the interaction, separate analyses were conducted on the first and second halves of the list. The visual suggestion of a generation effect in the second half of the list, but not in the first, was confirmed by this analysis. There was a highly reliable main effect of encoding condition for Positions 5–8, $F(1, 35) = 9.68$ ($MS_e = .03$) but no comparable effect for Positions 1–4, $F(1, 35) < 1$ ($MS_e = .02$).

Overall, the pattern confirms the major prediction of Experiment 3. There was no evidence of a negative generation effect in recall, even though that pattern was quite strong in Experiment 2. We have no firm explanation for the nature of the interaction, but it is interesting to note that the reconstruction data present a similar trend. The read advantage in Figure 5 is especially apparent at the end of the list, which is the mirror image of the free-recall results. One possibility, based on the reasoning of McDaniel, Riegler, and Waddill (1990), is that subjects tend to rely more on category information as the end of the list approaches. The more prior experience with category members in the list, the more informative category

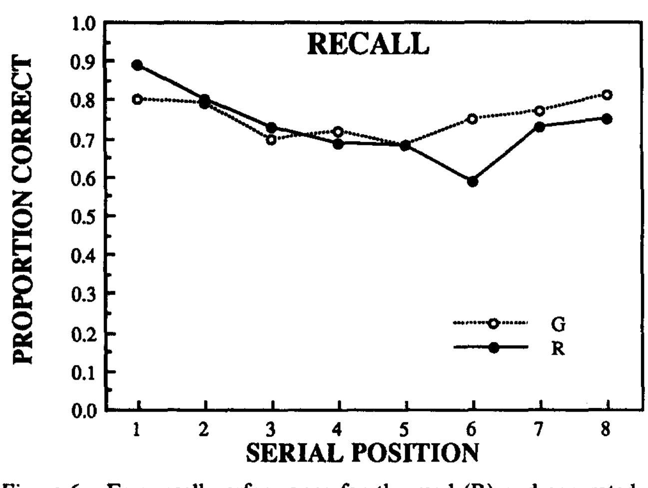

Figure 6. Free-recall performance for the read (R) and generated (G) lists in Experiment 3.

information becomes in solving the word fragment. Such categorical processing might then provide a further boost in individual-item memory, but at the expense of remembering the categorically irrelevant serial order of presentation. This reasoning again appeals to a tradeoff among the complex kinds of processing induced by generation. Increased categorical processing, which can be functionally relevant for generation, decreases any associative chaining process that might normally occur among adjacent items on the list.

To provide a measure of seriation, we again applied the Asch-Ebenholtz (1962) analysis to the free-recall data. Our prediction, that subjects would show less overall input–output correspondence with categorized lists, was confirmed. For generated lists, the average index of seriation was only 0.55 compared with 0.61 for the read lists. Both of these values were lower than those reported in Experiment 2, which when coupled with the higher overall levels of recall, strongly suggest that subjects were able to make use of an alternative organizational strategy to generate output candidates. In addition, as predicted, there was a statistically reliable difference in the average index for the read and generated lists, $t(35) = 3.16$. This last result substantiates the reconstruction data and provides a further indication that fragment completion interferes with the retention of serial order.

The results of Experiment 3 again present an illustration of dissociation, in which a single independent variable (generating vs. reading) affects two dependent variables in quite different ways. The presence of dissociative effects, although perhaps complicating the final interpretation, effectively rule out any simple-minded theoretical position on the effects of generation. It is quite clear, for example, that requiring subjects to complete fragments during a limited presentation period does not simply reduce overall retention of the presented material (as one might be tempted to conclude from the results of Experiment 2). Rather, the effects of generation are complex, and their revealed nature is likely to depend on one's choice of retention measure. Once again, this conclusion is neither new nor particularly surprising (see Begg et al., 1989; Schmidt & Cherry, 1989). However, similar reasoning needs to be applied to the analysis of a single dependent variable like free recall. Successful recall depends on many kinds of mnemonic information, any of which can be affected by an encoding procedure like generation. To ignore this fact is to invite a range of perplexing results and interpretative problems of the type currently surrounding the literature on the generation effect.

General Discussion

These experiments were undertaken to investigate the effects of generation on item and order retention. The item–order distinction has proven valuable in the analysis of immediate memory data but has yet to be translated effectively to the analysis of long-term memory phenomena (although see Nairne, 1990b). In the present case, generation produced dissociative effects on item and order retention, mimicking the kind of pattern found when categorical similarity is manipulated in immediate serial recall. Generation sometimes enhanced subjects' ability to remember the particular items that occurred on a trial but consistently interfered with the retention of serial order. Similarly, categorical relatedness increases the likelihood of item recovery in both a Brown–Peterson task and immediate serial recall but produces more transpositions in the orderings of those items at output (Crowder, 1979; Murdock & Vom Saal, 1967).

To draw on the short-term/long-term memory comparison a bit further, one might assume that generated items share common attributes, produced as a consequence of the generation act, that increase their net interitem similarity relative to read items. Increased feature overlap, according to an immediate memory perspective, should then make it easier to retrieve one generated item given another as a cue but harder to discriminate relative position within a list of generated items. One of us (Nairne, 1990a) has used a similar analysis to explain why continued articulatory suppression during list presentation impairs the retention of serial order information in immediate memory tasks. Perhaps more simply, it is possible that generation merely disrupts the normal associative chaining process that underlies the retention of serial order (e.g., Shiffrin & Cook, 1978); to the extent that a subject relies on interitem associations during recall, impaired memory (relative to read items) would be the anticipated result.

In each of the three experiments, we have presented evidence that fragment completion during presentation hurts a subject's ability to remember the relative orderings of items in a list. This finding was consistent for both related and unrelated lists, in both the reconstruction data and our index of seriation in free recall. Furthermore, this finding cannot be explained away as a simple consequence of the use of a data-limited presentation procedure. First, our presentation times did not differ dramatically from those used in other studies of the generation effect (2.5 s separated the onset of each item); Nairne (1988) showed a generation effect, for example, when items were presented for less than a second. Second, any appeal to a data-limited presentation procedure would predict general impairment across retention measures; our measures of individual-item memory revealed that generation increased the likelihood of correct recognition and recall under some circumstances. The fact that generation impaired serial order retention, even under conditions in which subjects were expecting an order test, strongly suggests that the disruption will be widespread.

With respect to the design controversy currently surrounding the generation effect, the present data point toward a rather straightforward resolution. Recent studies have shown that when the retention measure is free recall, the generation effect emerges only when reading and generating occur in the same list (see Begg & Snider, 1987; Hirshman & Bjork, 1988; Slamecka & Katsaiti, 1987). Between-list manipulations of generation have produced either no generation advantage or even a slight read advantage; the free recall results of Experiment 2 confirm this general pattern. These results become less perplexing when one recognizes that free recall (perhaps to a greater extent than recognition) relies on many forms of mnemonic information (cf. Begg et al., 1989; Schmidt & Cherry, 1989). Successful free recall requires some kind of organizational structure (for generation of output candidates) as well as individual-item information for use in deciding whether output candidates occurred in the study episode (e.g., Hunt & Einstein, 1981). An encoding procedure that improves the latter kind of information will aid recall only if it does little to disrupt the former.

We have shown that generation can produce dissociative effects on these two kinds of memories. Evidence gathered in the 1960s (see Postman, 1972, for a comprehensive review) documented that subjects often rely on serial order retention as a way of organizing output in free recall (particularly when list items are unrelated). Because generation reduces the effectiveness of seriation, the beneficial effects of this encoding procedure (enhanced individual-item memory) are likely to be masked. Such an account easily explains why between-list manipulations of reading versus generating can produce generation advantages in recognition, which often does not depend on good organizational structure (see Slamecka & Graf, 1978), but not in recall. The account also explains why between-list generation effects emerge in free recall when lists are composed of related items (see McDaniel et al., 1988). With related lists, inherent category structure affords an alternative output strategy, one that is not critically tied to serial order (see the results of Experiment 3). Finally, when reading and generating are conducted in the same list, it is reasonable to anticipate a similar disruption in serial order retention as a function of generation; however, when reading and generating are randomly mixed in a list, disruption in the ordering of one kind of stimulus will naturally affect retention of the other. Thus, any loss in the ability to recall by serial order will affect the list as a whole, unmasking a relative generation advantage.

Although this account provides a relatively complete explanation for a range of perplexing results, we recognize that our reasoning has a correlational flavor at this point. Just because generation disrupts serial order retention and the effectiveness of seriation in free recall does not mean that all instances of the presence and absence of generation effects have been explained as a result. As noted by virtually everyone, the effects of generation will depend on the generation procedure used and on the retention test adopted. There are probably situations, for example, in which the generation task will induce other complex forms of processing (cue–target relational processing, whole-list processing, etc.) that we have not addressed here. We believe that our empirical conclusions are firm and hope that at least some aspects of the literature have been simplified as a result.

Finally, we feel that the present experiments provide another clear example of the utility of the generation manipulation as a tool for the investigation of critical memory issues. In much the same way that manipulations of categorical structure have provided insight into organizational processes in recall and recognition, studies of the generation effect have enhanced our understanding of a number of memory issues, including transfer-appropriate processing, reality monitoring, and the influence of experimental design in mnemonic processing (among other examples). In the present instance, generation has been shown to be an effective tool for dissociating the retention of item and order information. Although long considered of importance in immediate retention, the item–

order distinction has yet to be translated effectively into the long-term memory arena. These experiments have documented the importance of the distinction in long-term retention and, as a result, should play an important role in furthering its study.

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.

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 Monograph, 80,* 1–45.

Begg, I., & Snider, A. (1987). The generation effect: Evidence for generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 553–563.

Begg, I., Snider, A., Foley, F., & Goddard, R. (1989). The generation effect is no artifact: Generating makes words distinctive. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 977–989.

Bellezza, F. S., Cheesman, F. L., & Reddy, G. (1977). Organization and semantic elaboration in free recall. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 539–550.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior, 13,* 80–87.

Bousfield, W. A., & Abramczyk, R. R. (1966). Sequential ordering in repeated free recall as a function of the length of the stimulus word list. *Canadian Journal of Psychology, 20,* 427–434.

Crowder, R. G. (1979). Similarity and order in memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 13, pp. 319–353). San Diego, CA: Academic Press.

Drewnowski, A. (1980). Attributes and priorities in short-term recall: A new model of memory span. *Journal of Experimental Psychology: General, 109,* 208–250.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12,* 360–372.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 644–655.

Healy, A. F. (1982). Short-term memory for order information. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 191–238). San Diego, CA: Academic Press.

Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 484–494.

Horowitz, L. M. (1961). Free recall and ordering of trigrams. *Journal of Experimental Psychology, 62,* 51–57.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 497–514.

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22,* 485–508.

Jahnke, J. C. (1965). Primacy and recency effects in serial-position curves of immediate recall. *Journal of Experimental Psychology, 70,* 130–132.

Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory* (pp. 331–373). San Diego, CA: Academic Press.

Mandler, G. (1969). Input variables and output strategies in free recall of categorized lists. *American Journal of Psychology, 82,* 531–539.

Mandler, G. (1980). Recognition: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

McDaniel, M. A., Riegler, G. L., & Waddill, P. J. (1990). Generation effects in free recall: Further support for a three-factor theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16,* 789–798.

McDaniel, M. A., Waddill, P. J., & Einstein, G. O. (1988). A contextual account of the generation effect: A three factor theory. *Journal of Memory and Language, 27,* 521–536.

Murdock, B. B., Jr. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General, 105,* 191–216.

Murdock, B. B., Jr. (1983). A distributed memory model for serial-order information. *Psychological Review, 90,* 316–338.

Murdock, B. B., Jr., & Vom Saal, W. (1967). Transpositions in short-term memory. *Journal of Experimental Psychology, 74,* 137–143.

Nairne, J. S. (1988). The mnemonic value of perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 248–255.

Nairne, J. S. (1990a). A feature model of immediate memory. *Memory and Cognition, 18,* 251–269.

Nairne, J. S. (1990b). Similarity and long-term memory for order. *Journal of Memory and Language, 29,* 733–746.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph, 76*(1, Pt. 2).

Postman, L. (1972). A pragmatic view of organization theory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 3–48). San Diego, CA: Academic Press.

Roediger, H. L., III. (1984). Does current evidence from dissociation experiments favor the episodic/semantic distinction? *The Behavioral and Brain Sciences, 7,* 252–254.

Schmidt, S. R., & Cherry, K. (1989). The negative generation effect: Delineation of a phenomenon. *Memory & Cognition, 17,* 359–369.

Shiffrin, R. M., & Cook, J. R. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior, 17,* 189–218.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604.

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language, 26,* 589–607.

Received May 25, 1990
Revision received November 2, 1990
Accepted December 17, 1990 ■