Item-specific processing and the generation effect: Support for a distinctiveness account

JOHN M. GARDINER and JAMES A. HAMPTON
The City University, London, England

Self-generated items are often better recalled than items that are read. Results of two experiments showed that this generation effect was greatly reduced or eliminated by an additional orienting task that was assumed to enhance item-specific processing in incidental learning. These results were predicted by a distinctiveness hypothesis according to which the generate task itself enhances item-specific processing. But the results are not inconsistent with certain other hypotheses that have been suggested to explain generation effects.

Memory performance has often been found to be superior for items that were self-generated at study rather than read (e.g., Donaldson & Bass, 1980; Gardiner & Hampton, 1985; Glisky & Rabinowitz, 1985; Graf, 1980; Jacoby, 1983; Nairne, Pusen, & Widner, 1985; Payne, Neely, & Burns, 1986; Slamecka & Graf, 1978). The attention attracted by these generation effects in recent years has led to a situation where it has become difficult to perceive the general significance of all the evidence now at hand. It has become clear, however, that no single one-factor hypothesis is going to prove capable of accounting adequately for this evidence. Indeed the evidence is broadly consistent with the involvement of at least three different factors in generation effects: (a) a rather general, conceptually-driven or semantic processing component; (b) a specific data-driven or surface processing component; and (c) the appropriateness of the test vis-à-vis study conditions.

Evidence for a general, conceptually-driven or semantic processing component has been described, for example, by Gardiner & Hampton (1985), Jacoby (1983), McElroy and Slamecka (1982), Payne et al. (1986), and Roediger and Blaxton (1987). This evidence includes finding no generation effects for pronounceable letter strings not represented as words in the internal lexicon (e.g., McElroy & Slamecka, 1982), or for items not represented as familiar integrated concepts in semantic memory (e.g., Gardiner & Hampton, 1985).

AMERICAN JOURNAL OF PSYCHOLOGY
Winter 1988, Vol. 101, No. 4, pp. 495–504

Glisky and Rabinowitz (1985) provided evidence for an additional, specific surface-processing component by showing that generation effects in recognition memory were enhanced if subjects were required at test to repeat an *identical* generate task to that used at study. Evidence for the role of test appropriateness comes from studies by Nairne and Widner (1987) showing that when test conditions were appropriate, in that they tapped or matched generate operations at study, there were generation effects for nonwords (see, too, Johns & Swanson, 1988). Further evidence for the role of test appropriateness comes from studies showing that priming effects in perceptual identification and word-fragment completion, which tend to depend more on data-driven than on conceptually driven processing, were greater for read than for generate items (Jacoby, 1983; Roediger & Blaxton, 1987; but see also Gardiner, in press).

The present article is concerned with the nature of the conceptually driven or semantic processing component of generation effects. More particularly, it describes some tests of a distinctiveness hypothesis that we had proposed previously (Gardiner & Hampton, 1985). This hypothesis was based on the conjecture that the evidence then available indicated that generation effects were due to item-specific rather than relational processing, in the sense that these two sorts of processing have been defined and evidenced as explanatory principles by Hunt and his colleagues (see, e.g., Einstein & Hunt, 1980; Hunt & Einstein, 1981; Hunt & Seta, 1984). In this distinction, item-specific processing corresponds with levels-of-processing, and relational processing corresponds with organizational coding. It is important to note that relational processing is not used here to refer to individual stimulus-response relations but to organizational relations among list items as a whole.[1] Individual item information is assumed to be important in delineating items in retrieval, that is, in enhancing their distinctiveness. In this account, relational and item-specific processing are used as explanatory principles, and do not refer to any specific mechanisms (Einstein & Hunt, 1980). The distinctiveness hypothesis, then, is simply the hypothesis that the generate task enhances item-specific processing, thereby facilitating subsequent discriminability in retrieval. If this hypothesis is correct, an orienting task that similarly enhances item-specific processing should significantly reduce, if not eliminate, the generation effect.

Support for the view that self-generation of words (for example by the replacement of missing letters) increases item-specific processing comes from a study by McDaniel, Einstein, Dunay, and Cobb (1986), who also used the conceptual framework of item-specific versus relational processing to derive predictions about generation effects. In

their task, subjects recalled one of two different passages of prose that were selected to contain either largely item-specific or relational information. Among their results, they showed that as predicted from the explanatory framework of Hunt and his colleagues, a generation task, which they assumed increased item-specific processing, resulted in an improvement in recall for only the relational information text. Thus, for the other text, where item-specific information was already strongly encoded, there was little additional advantage gained from generating items.

The two experiments reported here followed a plan similar to that of McDaniel et al. (1986). Assuming that the generation effect is at least in part due to enhanced item-specific processing, we predicted that an orienting task that provided such processing for both read and generate items, immediately after they have been presented, should reduce or even eliminate the generation effect.

EXPERIMENTS 1a, 1b

Experiments 1a and 1b were independently run replications of the same set of experimental conditions. Subjects in each replication were shown a list of members of different categories, half of which they had to generate, half of which they had to read. In addition, they were given incidental learning instructions and one of two orienting tasks of the sort that have been established by Hunt and his associates in earlier studies to entail item-specific and relational processing, respectively (e.g., Einstein & Hunt, 1980). The orienting task which drew attention to item-specific information was to rate the typicality of each individual category member as a member of its category. The other task, which reflected the obvious relational characteristics of the list, was to sort the items together into the appropriate category groups. If the generate task involves item-specific processing, then the generation effect should be significantly reduced, if not eliminated, following the typicality rating task, as compared with the category sorting task.

METHOD

Subjects

There were 48 subjects, 24 in Experiment 1a and 24 in Experiment 1b, all undergraduate students at the City University, London. Within each replication, 12 subjects were assigned alternately by order of testing to one of two independent groups.

Design

The design was a 2 × 2 × 2 factorial with experiment (1a vs. 1b) and orienting task (sort vs. rate) as between-subjects factors and generate condition (generate vs. read) as a within-subjects factor. The study list was 36 words, 6 from each of six different categories (birds, clothing, fish, flowers, fruit, and sport), selected from the Hampton and M. M. Gardiner (1983) norms so as to have an intermediate range of typicality ratings. Each of the following pairs of examples is, respectively, the most typical and least typical category member selected: blackbird, woodpecker; trousers, slippers; herring, pilchard; daffodil, hyacinth; apple, blackcurrant; tennis, skiing. Subjects had to generate half the members of each category and read the other half; generate and read items were fully counterbalanced. Subjects were given incidental learning instructions appropriate to one or the other orienting task and, later, an unanticipated free-recall test.

Procedure

The study list was presented on a deck of cards on each of which the name of a category appeared together with one category member. In the generate condition, all the vowels of the category member were omitted, a procedure that guaranteed 100% successful generation. Subjects had to say aloud both the name of the category and the category member for both generate and read items. Order of presentation was randomized separately for each subject. Subjects were told that the point of the experiment was to investigate the judgments people made about word characteristics. Subjects given the sorting task had to place the cards in groups under the appropriate category names; these category names were also displayed on the table at which subjects were placed. Subjects given the rating task had to rate the typicality of each category member on a form provided for the purpose, using a 5-point rating scale (1 = *very typical;* 5 = *very atypical).* The rating instructions were similar to those used by Hampton and M. M. Gardiner (1983). Subjects were told:

> You have to decide whether each word is a good example of the category named. For instance, most people would say that churches are very typical examples of the category *buildings,* more typical than, say, telephone boxes, which some people would classify as very atypical examples. The above example also serves to illustrate the point that just because a specific word is more typical than another, it does not mean that it occurs more often in your experience than an atypical word. Telephone boxes are probably seen much more often than churches.

Both tasks were self-paced. Performance in these tasks was monitored to ensure compliance with instructions, but otherwise unscored. After the tasks had been completed, subjects were given 60 s of a filler task, identifying particular sequences of digits from a puzzle matrix, to reduce the recall of recency items. They were then asked to write down in any order they liked, all the category members they could recall.

RESULTS AND DISCUSSION

The principal results are summarized in the left-hand panel of Figure 1. The figure shows that the results of Experiments 1a and 1b were essentially similar. In each case there was a large generation effect following the sorting task, and the generation effect following the rating task was greatly reduced, largely because of a rise in the scores for read items. These observations are supported by the results of an ANOVA carried out on subjects' individual recall scores. Recall was significantly higher for self-generated items, $F(1, 44) = 43.73$, $p < .001$, and for items that had been rated for typicality, $F(1, 44) = 21.76$, $p < .001$. The interaction between orienting task and generate condition was significant, $F(1, 44) = 8.47$, $p < .01$. No other effect or interaction was significant (all Fs < 1).

The enhancement in recall of such obviously related material by an orienting task that focuses attention on item-specific information replicates, in part, results previously reported by Einstein and Hunt (1980). More important for present purposes, this enhancement of recall greatly reduced the generation effect, as predicted by the distinctiveness hypothesis. However, although greatly reduced, some advantage to generate items remains apparent, and there was a slight rise in the scores for generate items, too. Because the orienting tasks were self-paced, one cannot tell whether this remaining advantage is of consequence or just an artifact of differences in processing times—a possibility rendered likely because subjects took somewhat longer

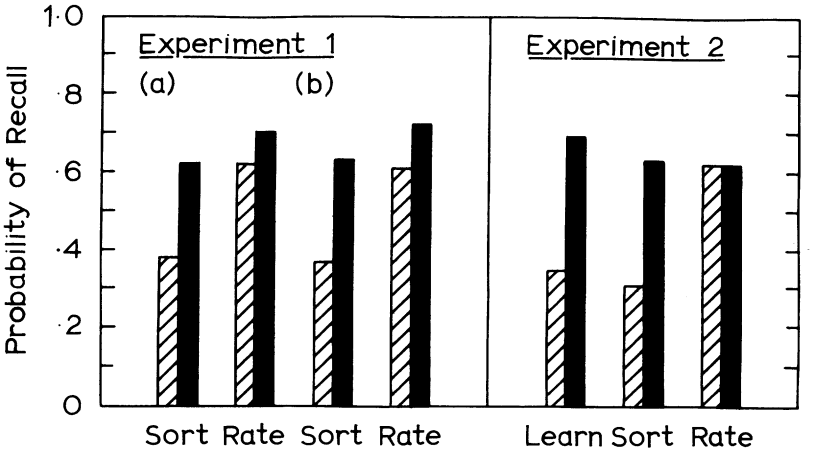

Figure 1. Probability of recall as a function of orienting task for generate (solid bars) and read (cross-hatched bars) items in Experiment 1 and Experiment 2

to judge typicality than to sort by category. Also, Experiment 1 omitted an intentional learning condition, and it would be of interest to compare the effects of the two incidental learning orienting tasks with those of instructions simply to learn the words. Experiment 2 was designed accordingly.

EXPERIMENT 2

Experiment 2 differed from Experiment 1 in two crucial respects: Presentation rate was paced by the experimenter; and there was a third group, which was given an intentional learning task. In view of evidence from organizational theorists that with highly related materials, sorting and learning instructions can be functionally equivalent (see, e.g., Mandler, 1967), it was expected that performance in the intentional learning group would be similar to that in the group given sorting instructions.

METHOD

Subjects

Subjects were 36 undergraduate students at the City University, London; 12 subjects were assigned by order of testing to each of three independent groups. None had participated in Experiment 1.

Design and procedure

The design was a 3 × 2 factorial with orienting task (sort, rate, or learn) as a between-subjects factor and generate condition (generate vs. read) as a within-subjects factor. In all other respects, save one, the design and procedure were identical with those of Experiment 1. In Experiment 2, items were presented by the experimenter at a rate of 7 s/card, a rate that allowed subjects sufficient time to make judgments of typicality.

RESULTS AND DISCUSSION

The principal results are summarized in the right-hand panel of Figure 1. The figure shows that performance following learning instructions was quite similar to that following the sorting task, and that in both cases there was a large generation effect. Following the rating task, however, the generation effect was eliminated altogether, entirely because of a rise in the scores for read items. These observations are supported by the results of an ANOVA carried out on subjects' individual recall scores. Recall was significantly higher for self-generated items, $F(1, 33) = 52.64$, $p < .001$. There was a significant difference

between orienting tasks, $F(2, 33) = 4.16$, $p < .025$. The interaction between orienting task and generate condition was significant, $F(2, 33) = 12.64$, $p < .001$. It therefore seems clear that learning and sorting tasks were functionally equivalent for both generate and read items, while typicality rating enhanced recall by raising the scores of only read items, to an extent that eliminated the generation effect. The explanation of the slight increase in performance for generate items under the rating condition in Experiment 1 in terms of an effect of self-pacing was supported by Experiment 2, where the generate items were recalled equally well under both orienting task conditions.

GENERAL DISCUSSION

The generation effect was greatly reduced or eliminated by an orienting task that was assumed to enhance item-specific processing in incidental learning. The elimination of the effect was due entirely to a rise in the scores for read items. It may thus be inferred, given our assumptions about the orienting tasks, that both generate and typicality rating tasks entail a similar form of item-specific processing. The results therefore provide good support for the distinctiveness hypothesis (Gardiner & Hampton, 1985), according to which the generate task itself enhances the encoding of item-specific information (Einstein & Hunt, 1980; Hunt & Einstein, 1981), thereby facilitating subsequent discriminability in retrieval.

These results are related to some earlier findings reported by Donaldson and Bass (1980), who tested the hypothesis that the generate task entails an implicit "adequacy check" of the link between each generated response word and its stimulus word. They had subjects perform an explicit adequacy check, judging how closely each response word was related to its stimulus word, and they found that this task significantly reduced the generation effect. Because in terms of the present distinction between relational and item-specific processing the adequacy check is another task that focuses attention on item-specific information, their results are entirely congruent with ours.

As McElroy and Slamecka (1982) have pointed out, however, the checking task used by Donaldson and Bass (1980) markedly raised the scores for generate as well as read items, which makes it difficult to argue convincingly that the two tasks—generating and checking—engaged the same processing. And there are other findings that are difficult to reconcile with the adequacy check hypothesis, including the failure to find generation effects for stimulus words (Payne et al., 1986; Slamecka & Graf, 1978) and the occurrence of generation effects in the absence of stimulus words (e.g., Gardiner & Hampton,

1985; Glisky & Rabinowitz, 1985; Nairne et al., 1985; but see also McElroy, 1987; Rabinowitz & Craik, 1986). Because these findings are quite consistent with the distinctiveness hypothesis, the latter is more general, and on those grounds to be preferred.

Of course, the logic of the present research does not permit the discounting of alternative interpretations that are based on other assumptions. Perhaps the most obvious alternative has recently been suggested in a provocative paper by Slamecka and Katsaiti (1987; see also Begg & Snider, 1987). On discovering that the generation effect in free recall could be obtained only when generate and read items were presented together in a mixed list, and on further showing that forcing equal rehearsal of each item eliminated the generation effect even in a mixed list, Slamecka and Katsaiti proposed that the standard generation effect in recall is simply an artifact due to selective displaced rehearsal of generate items. However, as we have argued elsewhere (Gardiner, Gregg, & Hampton, in press), their account cannot be a general explanation of generation effects, and indeed Slamecka and Katsaiti themselves acknowledged as much. Also, Hirshman and Bjork (in press) found substantial generation effects in cued recall, if not free recall, with an unmixed list (between-subjects) design. Because only six categories were used in the present experiments, functionally the tests correspond with cued recall rather than free recall, in that the category names undoubtedly served as implicit retrieval cues. Moreover, a selective displaced rehearsal account of our results would have to assume that despite incidental learning instructions, subjects selectively rehearsed generate items and that this selective rehearsal was eliminated by the rating task but not by the sorting task. Though conceivable, such an interpretation seems rather strained.

However, the finding of generation effects in free recall with mixed but not unmixed list presentation is itself not inconsistent with an account in terms of the distinctiveness principle, and it is also possible that enforcing equal rehearsal of read items and generate items in a mixed list may increase item-specific processing, and so enhance distinctiveness. Perhaps further research may reveal that these two accounts are not so incompatible as they might now seem.

Notes

Experiment 1a was carried out in collaboration with Amanda Tomes as part of an undergraduate project; we are also grateful to Rosalind Java and Christine Turgoose for experimental assistance. Offprint requests should be addressed to John M. Gardiner, Memory & Cognition Research Group, The City University, Northampton Square, London EC1V 0HB, England, U.K.

Received for publication October 21, 1987; revision received December 17, 1987.

1. Some studies that have not embraced this distinction exactly as proposed have used the term relational processing to refer to the encoding of pairwise stimulus-response relations (e.g., Begg & Snider, 1987; Donaldson & Bass, 1980; Hirshman & Bjork, in press).

References

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

Donaldson, W., & Bass, M. (1980). Relational information and memory for problem solutions. *Journal of Verbal Learning and 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 and Memory, 6,* 588–598.

Gardiner, J. M. (in press). Generation and priming effects in word-fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Gardiner, J. M., Gregg, V. M., & Hampton, J. A. (in press). Word frequency and generation effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Gardiner, J. M., & Hampton, J. A. (1985). Semantic memory and the generation effect: Some tests of the lexical activation hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 732–741.

Glisky, E. L., & Rabinowitz, J. C. (1985). Enhancing the generation effect through repetition of operations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 193–205.

Graf, P. (1980). Two consequences of generating: Increased inter- and intraword organization of sentences. *Journal of Verbal Learning and Verbal Behavior, 19,* 316–327.

Hampton, J. A., & Gardiner, M. M. (1983). Measures of internal category structure: A correlational analysis of normative data. *British Journal of Psychology, 74,* 491–516.

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

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and 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, and Cognition, 10,* 454–464.

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

Johns, E. E., & Swanson, L. G. (1988). The generation effect with nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 180–190.

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 327–372). New York: Academic Press.

McDaniel, M. A., Einstein, G. O., Dunay, P. K., & Cobb, R. E. (1986). Encoding difficulty and memory: Toward a unifying theory. *Journal of Memory and Language, 25,* 645–656.

McElroy, L. A. (1987). The generation effect with homographs: Evidence for post-generation processing. *Memory & Cognition, 15,* 148–153.

McElroy, L. A., & Slamecka, N. J. (1982). Memorial consequences of generating nonwords: Implications for theories of the generation effect. *Journal of Verbal Learning and Verbal Behavior, 21,* 249–259.

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, and Cognition, 13,* 164–171.

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.

Rabinowitz, J. C., & Craik, F. I. M. (1986). Specific enhancement effects associated with word generation. *Journal of Memory and Language, 25,* 226–237.

Roediger, H. L., III, & Blaxton, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and cognitive processes: The Ebbinghaus centennial conference.* Hillsdale, NJ: Erlbaum.

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.