The generation effect with homographs: Evidence for postgeneration processing

LORI A. MCELROY

University of Toronto, Toronto, Ontario, Canada

The generation effect is the phenomenon in which words are remembered better when generated than when read. These experiments test the possibility that at least one consequence of generating is enhanced semantic processing. Homographs were used as targets, presented with rhymes in Experiment 1 so as not to bias meaning, and with synonyms in Experiment 2 to bias one meaning of each homograph. In both experiments, extralist synonym cues were provided at recall. In Experiment 1 a generation effect was obtained when the retrieval cues biased the dominant meaning of the homograph (determined from free association norms), whereas in Experiment 2 a generation effect was found when the retrieval cues biased the same meaning that was biased during study. In neither experiment was a generation effect obtained with retrieval cues that biased the other meaning of each homograph. These results indicate that the generation effect is dependent upon the compatibility of the semantic processing conducted at study and test. Since it is impossible to process the meaning of a homograph when generating it from a rhyme cue, the meaning of the homograph could only have been processed after the word had been generated. The finding in Experiment 1 that a generation effect was obtained with rhymes when semantic retrieval cues were provided demonstrates that the enhancement properties associated with generation are not restricted to the information used to guide the generation process. This finding also indicates that one locus of the generation effect is in the processing that occurs after the word has been generated.

The generation effect refers to the memory advantage resulting from active participation in the learning phase (Slamecka & Graf, 1978). In the typical generation task, the response member of a paired associate is generated using a semantic, phonemic, or structural rule relating the response word to the study cue. Performance for the generate condition is compared with that of the more passive condition in which the same material is only read. Superior retention of generated words has been demonstrated on free recall, intralist-cued recall, and item recognition tests (Donaldson & Bass, 1980; Gardiner & Arthur, 1982; Gardiner & Hampton, 1985; Jacoby, 1978; Slamecka & Graf, 1978). Similar results have been obtained when an entire sentence is constructed by arranging a list of words according to a specified grammar (Graf, 1980, 1982) and when words are generated from word fragments rather than from associates (Glisky & Rabinowitz, 1985).

Although the effect is robust, it has been shown to depend upon the type of materials studied. While the phenomenon is readily obtained with relatively common

words, it has not been found with nonsense words on such standard retention tests as free recall and word recognition (Gardiner & Hampton, 1985; McElroy & Slamecka, 1982; Nairne, Puse, & Widner, 1985; Payne, Neely, & Burns, 1986). A generation effect has been obtained with nonsense words only on very specialized tests (Nairne & Widner, in press). For these experiments, the words and nonsense words were generated through the transposition of two underlined letters in a cue. For instance, the nonword PERZIK would be produced from the cue ZERPIK, and the word HEAVEN would be produced from the cue VEAHEN. A generation effect was obtained for both words and nonwords on a test in which subjects had to choose PERZIK from PERZIK and PERZIK and HEAVEN from HEAVEN and HEAVEN, and on a test in which the words and nonwords had to be generated at test. On the other hand, a generation effect emerged with words, but not with nonsense words, on a standard item recognition test.

In a similar vein, Glisky and Rabinowitz (1985) recently demonstrated that the magnitude of the generation effect obtained with words can be enhanced by having the words generated at test, rather than read. They concluded that this enhancement is due to the repetition of the specific operations used to generate the word. This explanation could also be applied to Nairne and Widner's (in press) nonword data, since a generation effect emerged with nonwords only when the encoding operations were reinstated at test. However, as Glisky and Rabinowitz (1985) suggested, a repetition-of-operations view cannot provide a

This research was supported by National Science and Engineering Research Council of Canada Operating Grant A7663 to N. J. Slamecka, and by an Ontario Graduate Scholarship to the author. Experiment 2 was part of a thesis submitted to the University of Toronto in partial fulfillment of the requirements for the master's degree. The comments on earlier drafts of this manuscript by F. I. M. Craik, Ron Fisher, Jim Nairne, Barry Stein, and especially by Norman J. Slamecka are gratefully acknowledged. Requests for reprints should be addressed to Lori McElroy, Department of Psychology, University of Toronto, Toronto, Ontario, Canada M5S 1A1.

complete explanation of the generation effect, since a substantial generation effect is readily obtained with words even when the test does not reinstate the generation operations.

What, then, gives generated words an advantage that generated nonwords do not have? One difference between words and nonsense words is that words have meaning, whereas nonsense words do not, at least not to the extent that words do. The fact that words are amenable to semantic processing could explain why the generation effect occurs with words even on tests in which the generation operations are not reinstated. The experiments reported here examine the involvement of semantic encoding in the generation effect.

Since a generation effect can be obtained with phonemic and structural rules (Gardiner & Arthurs, 1982; Gardiner & Hampton, 1985; Glisky & Rabinowitz, 1985; Nairne et al., 1985; Slamecka & Graf, 1978), it is clear that the generation task itself need not involve semantic processing. It is always possible, however, that semantic processing occurs *after* the item has been generated. Perhaps generating, because of the active involvement required of the subject, induces more elaboration of the generated word than does reading, when such elaboration is possible. This notion was tested by Rabinowitz and Craik (1986), who reasoned that if a generation effect enhances the encoding of semantic information even for rhymes, then a generation effect should be obtained with semantic retrieval cues. They failed to find a generation effect with rhymes when semantic associates were provided as retrieval cues, although they did obtain a generation effect with rhyme cues. They concluded that generating only enhances the information that was used to guide the generation process. However, there is an alternate interpretation for their results. Since most common words generally have more than one semantic interpretation, Rabinowitz and Craik may have failed to obtain a generation effect with semantic cues simply because the semantic cues they used did not bias the same interpretation as was encoded during study. Experiment 1 was designed as a further test of whether semantic processing could occur after the word has been generated.

EXPERIMENT 1

To prevent semantic processing during generation, homographs were read or generated as rhymes. The homographs were presented with cues that differed in spelling from the target in their first letter only, and were unrelated to either meaning of the target (e.g., *luck duck*). At recall, extralist cues were provided that were related to one of the two meanings of each homograph. For half of the targets the cue biased the most common or dominant interpretation of the homograph, and for the other half the less common interpretation was biased. The dominant interpretation was the one most subjects gave for the homograph when presented without context.

Lexical decision research has shown that when the encoding context is neutral with respect to meaning, only the dominant meaning of the homograph is encoded (Simpson, 1981). Therefore, since rhyme cues do not bias any meaning, only the dominant sense of the homograph should be encoded, to whatever extent semantic processing occurs at all. Because it is impossible to know the meaning of the homograph until after it has been generated, if semantic processing is to occur for generated homographs, it must occur in the postgeneration phase. If generation does not induce postgeneration semantic processing, then the dominant meaning of the homograph should be encoded to the same extent for read and generate items, and no generation effect should be found regardless of which meaning is biased at test. If, on the other hand, generation leads to greater semantic processing than does reading, then a generation effect should be obtained at least with cues biasing the dominant meaning of the homograph. If this enhanced semantic processing is restricted to the dominant meaning, then the effect should be obtained only when the retrieval cue biases the dominant meaning. If the semantic processing that generate items receive causes all meanings of the homograph to be activated, then a generation effect should be obtained with both types of semantic cues.

Method

Subjects and Design. Twelve students from a third-year laboratory course at the University of Toronto participated as part of their course requirements. A within-subjects design was used with task (read or generate), meaning biased at retrieval (dominant or nondominant), and type of homograph (balanced or polarized) as the three factors.

Materials. Forty-six cue–target pairs were used in the input list, including 14 nonhomographs to disguise the homographic nature of the critical targets. The 32 homophonic homographs were selected from word association norms (Cramer, 1970; Kausler & Kollasch, 1970; Nelson, McEvoy, Walling, & Wheeler, 1980) that tabulated the probability of each meaning being activated when the homograph was presented without context. The meaning with a greater word association probability was classified as dominant. The homographs were subdivided into two groups, polarized or balanced, on the basis of the relative frequency of their two meanings. For polarized homographs, the word association probability for the dominant meaning was greater than .60, whereas the association probability for the nondominant meaning was less than .30. For balanced homographs, the two meanings were more equally probable, with both word association probabilities falling between .60 and .30. This classification was included as a precautionary measure because the encoding of balanced homographs, when presented without context, has been found to be less stable over time than the encoding of polarized homographs (Winograd & Geis, 1974). For recall, every homograph had one cue to bias each of its two meanings, and every nonhomograph had one cue biasing its only meaning. Across all subjects, each homograph was presented equally often in the read and generate conditions and was tested equally often by dominant and nondominant cues.

Study booklets were constructed with one cue–target pair typed on each page. For read items both cue and target were fully present (e.g., *luck duck*). Although the cue was unaltered for generate items, the internal vowels of the target were omitted and replaced by dashes

(e.g., *luck d-ck*). The items were presented in random order with the restriction that no more than three homographs appear in succession. In addition, three nonhomographs appeared at the beginning and end of the list to minimize primacy and recency effects. Read and generate items were randomly dispersed throughout the list. The booklets for cued recall contained one cue on each page. The cues were arranged randomly with respect to the original input order of targets, but again the restriction was made that no more than three homographs were cued in succession.

Procedure. Subjects were tested in one group in a classroom setting. So that generation accuracy could be monitored, subjects were required to write in their booklets the missing letters of generated words. The nature of the cue-target relationship was explained, with instructions to fill in incomplete words by substitution of the vowels from the cue. Subjects were given practice with the input phase and then told to expect a recall test following the input phase. Input was paced at 8 sec per item to allow sufficient time to fill in the missing letters. The recall phase followed immediately thereafter. With the aid of examples, subjects were informed that the cues were related to the target on the basis of meaning, and subjects were instructed to write beside each cue the target it reminded them of. Recall was paced at 8 sec per cue.

Results and Discussion

Since there were no input errors, recall probabilities are based on the total number of each type of item studied at input. The means for each combination of cue bias and task are presented in Table 1, separately for balanced and polarized homographs. From the table we can see that recall was better overall when the retrieval cues biased the dominant rather than the nondominant meaning of the homographs, and that a generation effect emerged only when the dominant meaning was biased at test.

Analysis of variance confirmed these observations. Unless otherwise specified, the alpha level was set at .05 for all analyses. In addition to significant main effects of task [$F(1,11) = 7.68$, $MSe = .019$] and bias [$F(1,11) = 12.48$, $MSe = .023$], there was a significant interaction of the two [$F(1,11) = 6.08$, $MSe = .013$]. Further analysis of the interaction indicated that a generation effect emerged when cues biased the dominant meaning [$t(11) = 3.27$], but not when the nondominant meaning was cued ($|t| < 1$).

Although the overall difference between the two types of homographs was not reliable [$F(1,11) = 2.2$, $MSe = .05$], the interaction of type of homograph and retrieval bias approached significance [$F(1,11) = 3.33$, $MSe = .033$, $p < .10$]. This interaction was due to the fact that dominant meaning cues were more effective for polarized homographs than for balanced homographs [$t(11) = 2.01$, $p < .10$], whereas there was no differ-

ence between the two types of homographs with nondominant cues ($|t| < 1$). However, type of homograph did not interact with task or with the bias $\times$ task interaction (both $F_s < 1$); this is particularly important in demonstrating that the level of read performance does not determine whether a generation effect is obtained. For polarized homographs recall of read items is higher to dominant cues than to nondominant cues [$t(11) = 2.1$, $p < .10$], whereas for balanced homographs recall of read items to dominant and nondominant cues was equivalent ($|t| < 1$). Nonetheless, the generation effect was reliable for both polarized and balanced homographs when the dominant meaning was biased at retrieval [$t(11) = 2.49$ and $t(11) = 2.31$, respectively]. When the nondominant meaning was biased at retrieval, the difference between read and generate was not reliable for either type of homograph (both $|t|s < 1$).

These results indicate that, even for rhymes, there is a difference in the degree to which meaning is encoded for read and generate items. The dominant meaning of homographs had a greater probability of being encoded for generate than read rhymes, as evidenced by the greater effectiveness of dominant cues for generated homographs. The meaning of a generate item could not have been encoded until after the item had been generated. These results clearly demonstrate that when the generation task does not involve semantic processing, semantic processing is involved in the postgeneration phase. They also indicate that when meaning is not biased by the encoding context, it is the subject's preexperimental experience with the language that determines the meaning that is enhanced by generating. The failure to obtain a generation effect when the nondominant meaning was biased is consistent with Rabinowitz and Craik's (1986) results. However, the fact that a generation effect can be obtained with rhymes when the right semantic cue is provided indicates that their conclusions were wrong. The consequences of generating are not restricted to the information used to guide the generation process.

It has generally been assumed that since the encoding context must be used to generate the target, the integration of context and target must be greater for generate than for read items (Donaldson & Bass, 1980; Graf, 1980, 1982; Jacoby, 1978; Slamecka & Graf, 1978). This integration presumably results in greater elaboration of relational information for generate items. Both Graf (1980, 1982) and Donaldson and Bass (1980) have demonstrated that there is greater processing of the cue-target relationship for generate items when meaning forms the basis of the relationship. Experiment 2 examined the impact of the cue-target relationship on the semantic processing that read and generate items receive.

EXPERIMENT 2

The general procedure of Experiment 2 consisted of a study phase in which homographs were read or generated in the presence of a study cue that biased either the

Table 1
Cued Recall Probabilities (and Standard Deviations) for Polarized and Balanced Homographs, as a Function of Encoding Task and Meaning Cued (Dominant or Nondominant) for Experiment 1

Test Bias	Type of Homograph			
	Polarized		Balanced	
	Read	Generate	Read	Generate
Dominant	.21 (.18)	.35 (.20)	.08 (.12)	.21 (.23)
Nondominant	.08 (.12)	.13 (.13)	.10 (.20)	.10 (.23)

homograph's dominant or nondominant meaning. Recall of the homographs was then tested in the presence of extralist cues related either to the meaning biased at input or to the alternate meaning. The study cue used to bias one of the homograph's two meanings should determine the semantic processing that occurs during generation. If semantic processing is completely determined by the encoding context, then a generation effect should be found only with cues that have the same semantic relationship to the target as the study cue. If, as was found in Experiment 1, one's preexperimental experience with the language also influences the semantic processing that generate items receive, then a generation effect should also be found when the nondominant meaning is biased during study and the dominant meaning is biased at retrieval, although the magnitude of the effect may be reduced. No generation effect is expected when the dominant meaning is biased during study and the nondominant meaning is biased at test.

Method

Subjects and Design. There were 16 subjects; 3 were paid volunteers, and 13 were University of Toronto introductory psychology students earning bonus course credits for their participation. The design had three factors: meaning biased during input (dominant or nondominant) and encoding task (read or generate) as within-subjects factors, and meaning biased on the initial cued recall test (same or different from meaning biased at input) as a between-subjects factor. In addition, all subjects were given a second cued recall test for which the original study cues were provided.

Materials. The study list consisted of 24 homograph and 18 nonhomograph targets. Nonhomographs were again included to ensure that the homographic nature of the critical targets was not apparent. Targets were related to their cues on the basis of meaning, as synonyms or associates. Each homograph had four cues, two to bias its dominant meaning and two to bias its nondominant meaning, whereas nonhomographs had only two cues, both biasing the same meaning. Across all subjects, each cue occurred equally often as an input and extralist cue. Each cue-target pair was typed on 3 × 5 in. file cards, once in the read and once in the generate format.

The input list consisted of 52 cue-target pairs presented in a random order with the restriction that no more than three homographs appear in succession. To minimize primacy and recency effects, three nonhomograph buffers were added at both the beginning and end of the list. Half of each of the two types of targets were read and half were generated, with read and generate items dispersed randomly throughout the list.

For the cued recall tests, booklets were devised that resembled those used in Experiment 1, with one cue appearing on each page. The cues were arranged randomly with respect to the original input order of targets, but again the restriction was made that no more than three homographs were cued in succession. Two different restricted random orders were used for the two successive tests.

Procedure. The study procedure was similar to that used in Experiment 1. After receiving instructions, subjects practiced the input phase and were told to expect a recall test of the target words. Input was paced at 4 sec, and each cue and target were spoken aloud.

Instructions for the first cued recall test were given immediately after the study phase. For the first test, half the subjects had cues eliciting the same meaning (same-meaning group) that was biased at input, and the rest had cues eliciting the different meaning (different-meaning group). Subjects were told that "the cues are meaningfully related to the target, but they were not presented with the targets when you studied the list." The testing procedure and

pacing were identical to those in Experiment 1, except that a test with input cues was administered following the extralist cued recall test. Both recall tests were paced at a rate of 8 sec per cue. To illustrate, the pair *coal mine* was studied by both groups, designated by their first test experience as the same- and different-meaning groups. On the first test the retrieval cue *gold* was provided for the different-meaning group; for the second test both groups received *coal* as the retrieval cue.

Results and Discussion

During the study phase the different-meaning group made errors on 3% of read homographs and 24% of generate homographs, whereas the same-meaning group made no errors on read homographs but made errors on 13% of generate homographs. Although the different-meaning group made about twice as many errors as did the same-meaning group, the pattern of errors for both groups was similar in that more errors occurred on generate than on read items. Those items on which errors were made were omitted from the scoring. A stringent scoring criterion was used throughout; credit was given only for targets written beside the appropriate cue.

The results of the first test, in which extralist same- or different-meaning cues were provided, will be discussed first. These recall probabilities, as a function of encoding bias and task, are presented in Table 2. It is apparent from the table that the group receiving same-meaning cues showed a substantial generation effect, whereas there was no generation effect for the group receiving different-meaning cues.

These observations were supported by an analysis of variance with test group (same- or different-meaning cue) as a between-subjects factor and task (read or generate) and encoding bias as within-subjects factors. The main effect of test group [$F(1,14) = 21.54$, $MSe = .025$] and the interaction of test group and task [$F(1,14) = 7.87$, $MSe = .022$] were significant. Additional analysis revealed a significant generation effect for the group receiving same-meaning cues at retrieval [$t(14) = 4.55$], but the difference between read and generate for the different-meaning group was not significant [$t(14) = -1.07$]. Test group also interacted with encoding bias [$F(1,14) = 9.20$, $MSe = .025$]. This interaction reflects the fact that for

Table 2
Recall Probabilities (and Standard Deviations) for Experiment 2, as a Function of Encoding Bias (Dominant or Nondominant Meaning), Task (Read or Generate), and Test Group (Same or Different Meaning Cues on Test 1), for Test 1 and Test 2

Test Group	Encoding Bias			
	Dominant		Nondominant	
	Read	Generate	Read	Generate
Test 1: Recall to Extralist Cues				
Same Meaning	.25 (.20)	.48 (.14)	.19 (.14)	.29 (.20)
Different Meaning	.10 (.15)	.03 (.07)	.18 (.14)	.17 (.15)
Test 2: Recall to Study Cues				
Same Meaning	.31 (.26)	.59 (.20)	.31 (.14)	.48 (.17)
Different Meaning	.32 (.23)	.58 (.24)	.27 (.21)	.40 (.30)

the group receiving same-meaning cues, recall was better when the dominant rather than the nondominant meaning of the homograph had been biased at encoding [$t(14) = 3.30$], whereas the reverse was found for the group receiving the different-meaning cues [$t(14) = -2.79$]. In other words, for both groups recall was better when the retrieval cue biased the dominant rather than the nondominant meaning of the homograph, regardless of which meaning was biased at study. Nonetheless, the generation effect emerged only when the retrieval cue was related to the meaning that was biased at encoding, regardless of whether this was the dominant or nondominant meaning; both the interaction of bias and task and the interaction of test $\times$ bias $\times$ task were nonsignificant.

One-tailed t tests were performed to further demonstrate that the interaction of test group and task was not due to the level of recall in the read condition. When the dominant meaning was biased at encoding, a generation effect was obtained when the same (dominant) meaning was cued at test [$t(14) = 4.35$], but not when the different (nondominant) meaning was cued at test [$t(14) = -1.32$]. When the nondominant meaning was biased at encoding, a generation effect was still obtained when the same (nondominant) meaning was cued at test [$t(14) = 1.89$], but not when the different (dominant) meaning was cued at test ($|t| < 1$).

In Experiment 1, where the generation rule did not require semantic processing, the semantic processing that generated homographs received was determined by the subject's preexperimental experience with the language. The effect of preexperimental familiarity was also apparent in Experiment 2, in that retrieval cues biasing the dominant meaning were more effective than cues biasing the nondominant meaning, regardless of which meaning was biased at input. The occurrence of the generation effect, however, was dependent not on the meaning biased at test per se, but on the compatibility of the semantic processing performed during study and test. A generation effect emerged only when the test required the same semantic processing as that required at study. It should be noted that when the nondominant meaning was biased at encoding, recall for read items was virtually identical regardless of which meaning was biased at test. Despite the fact that the level of read performance was equal for the two groups, a generation effect occurred only with the group receiving the same-meaning cues. This indicates that the failure to obtain a generation effect is not dependent on the level of read performance. It also demonstrates that, as in Experiment 1, the semantic information that is enhanced by generating is specific to one meaning of the homograph.

Since the critical comparison was between subjects, it is possible that the different-meaning group, for some unknown reason, did not learn the generated items to the same extent as did the same-meaning group. That this concern is unfounded is readily apparent by examination of

the means associated with recall to study cues, which are provided in Table 2. When the same analysis was performed on these data, only a main effect of task emerged [$F(1,14) = 43.04$, $MSe = .016$]. When study cues were provided at retrieval, recall was always superior for generate items, and the performance of the two groups was indistinguishable.

GENERAL DISCUSSION

In both Experiments 1 and 2, a generation effect was found with homographs only when retrieval cues biased the same meaning as was encoded at study. This was true whether the encoded meaning was explicitly biased during input, as in Experiment 2, or determined by the subjects' preexperimental experience with these items, as in Experiment 1. Regardless of whether semantic processing was required by the encoding rule used, the requirement of generating enhanced the extent to which one specific meaning of the homographs was processed.

It may be that generating involves more extensive processing of an item's semantic memory attributes, including meaning, simply because generating is less automatic than is reading. It should be noted that this processing is beyond that merely required by the generation task. In particular, for rhyme generation such as *luck d*, no extensive processing is required to generate *duck*; a simple substitution of *d* for the *l* in *luck* will produce the required response. Yet the likelihood that a particular meaning of *duck* will be encoded is much greater if the word is generated than if it is read. When generating the rhyme *duck*, it is impossible to process its meaning until *after* the word has been generated. The fact that meaning is processed at all for rhymes suggests that the processing used to generate is not important as long as the product of generation can be processed semantically. What is done to arrive at the target, whether it is generated as a rhyme or as a synonym to a cue (Gardiner & Arthurs, 1982; Slamecka & Graf, 1978) or generated as a completion to a word fragment (Glisky & Rabinowitz, 1985) or generated by the transposition of letters (Gardiner & Hampton, 1985), may be irrelevant. The results of this research suggest that the failure to find a generation effect with nonwords (Gardiner & Hampton, 1985; McElroy & Slamecka, 1982; Nairne et al., 1985; Payne et al., 1986) on free recall and word recognition tests may be because enhanced semantic processing of nonwords is impossible.

Enhanced semantic processing has often been shown to yield a retention advantage (e.g., Craik & Tulving, 1975; Hyde & Jenkins, 1969). One possible reason for this advantage is that semantic processing may result in a more distinctive memory trace of the event, thus making the item easier to discriminate from other items in memory (Moscovitch & Craik, 1976; Stein, 1978). One consequence of generating may be that it enhances the en-

coding of only one specific meaning of the generated word, thereby conferring a retrieval advantage by making generated items more distinctive in memory.

REFERENCES

- CRAIK, F. I. M., & TULVING, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, **104**, 268-294.
- CRAMER, P. (1970). A study of homographs. In L. Postman & G. Keppel (Eds.), *Norms of word association* (pp. 361-382). New York: Academic Press.
- DONALDSON, W., & BASS, M. (1980). Relational information and memory for problem solutions. *Journal of Verbal Learning & Verbal Behavior*, **19**, 26-35.
- GARDINER, J. M., & ARTHURS, F. S. (1982). Encoding context and the generation effect in multitrial free-recall learning. *Canadian Journal of Psychology*, **36**, 527-531.
- 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, & 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, & Cognition*, **11**, 193-205.
- GRAF, P. (1980). Two consequences of generating: Increased inter- and intra-word organization of sentences. *Journal of Verbal Learning & Verbal Behavior*, **19**, 316-327.
- GRAF, P. (1982). The memorial consequences of generation and transformation. *Journal of Verbal Learning & Verbal Behavior*, **21**, 539-548.
- HYDE, T. S., & JENKINS, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, **82**, 472-481.
- 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.
- KAUSLER, D. H., & KOLLASCH, S. F. (1970). Word associations to homographs. *Journal of Verbal Learning & Verbal Behavior*, **9**, 444-449.
- MC ELROY, L. A., & SLAMECKA, N. J. (1982). Memorial consequences of generating nonwords: Implications for semantic-memory interpretations of the generation effect. *Journal of Verbal Learning & Verbal Behavior*, **21**, 249-259.
- MOSCOVITCH, M., & CRAIK, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding. *Journal of Verbal Learning & Verbal Behavior*, **15**, 447-458.
- NAIRNE, J. S., PUSEN, C., & WIDNER, R. L., JR. (1985). Representations in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition*, **13**, 183-191.
- NAIRNE, J. S., & WIDNER, R. L., JR. (in press). Generation effects with nonwords: The role of test appropriateness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.
- NELSON, D. L., McEVOY, C. M., WALLING, J. R., & WHEELER, J. W., JR. (1980). The University of South Florida homograph norms. *Behavior Research Methods & Instrumentation*, **12**, 16-37.
- 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 Experimental Psychology: Learning, Memory, & Cognition*, **25**, 226-237.
- SIMPSON, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning & Verbal Behavior*, **20**, 120-136.
- SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 592-604.
- STEIN, B. S. (1978). Depth of processing reexamined: The effects of the precision of encoding and test appropriateness. *Journal of Verbal Learning & Verbal Behavior*, **17**, 165-174.
- WINOGRAD, E., & GEIS, M. F. (1974). Semantic encoding and recognition memory: A test of encoding variability theory. *Journal of Experimental Psychology*, **102**, 1061-1068.

(Manuscript received June 12, 1985;
revision accepted for publication July 21, 1986.)