

# The generation effect: Further tests of the lexical activation hypothesis

DAVID G. PAYNE

*State University of New York at Binghamton, Binghamton, New York*

JAMES H. NEELY

*Purdue University, West Lafayette, Indiana*

and

DANIEL J. BURNS

*State University of New York at Binghamton, Binghamton, New York*

Three experiments compared recognition memory for word versus nonword responses when they had been either read or generated using a rhyme rule and either a word or nonword stimulus. That is, either the word *shop* or the nonword *thop* was generated from either the word *chop* or the nonword *phop*. In Experiment 1, the lexicality of the stimulus and response terms was manipulated orthogonally between subjects; in Experiments 2 and 3, it was manipulated orthogonally within subjects. In Experiment 3, subjects also made a lexical (word-nonword) decision about each response term after it had been read or generated. In all three experiments, memory performance was better for generated than read responses. This generation effect occurred only if the response term was a word, regardless of whether the stimulus term was a word or a nonword. These results are discussed in terms of the roles that lexical memory and response unitization play in the generation effect.

People are better able to recall or recognize words if they have previously produced or generated them than if they have merely read them. This so-called generation effect was first studied extensively by Slamecka and Graf (1978). In their generate condition, subjects used a specified rule (e.g., antonym) and a stimulus item followed by a single letter (e.g., HOT-C\_\_\_) to generate and say aloud the target response item (COLD); in their read condition, subjects also said aloud the response item (COLD), but both the stimulus and response items were presented intact (i.e., HOT-COLD). Subsequent tests showed better memory for the response items from the generate pairs than from the read pairs. This generation effect occurs for a wide range of materials and study and testing conditions (e.g., Anderson, Goldberg, & Hidde, 1971; Gardiner & Arthurs, 1982; Gardiner & Rowley, 1984; Glisky & Rabinowitz, 1985; Graf, 1980, 1982; Jacoby, 1978, 1983; Kane & Anderson, 1978; McFarland, Frey, & Rhodes, 1980; Slamecka & Fevreski, 1983; Stein & Bransford, 1979; Winnick & Daniel, 1970).

Several recent studies have demonstrated some of the boundary conditions for producing the generation effect (e.g., Donaldson & Bass, 1980; Gardiner & Hampton, 1985; Graf, 1980; McElroy & Slamecka, 1982; Nairne, Puse, & Widner, 1985). Of particular relevance to the present research is the finding that a generation effect does not occur for nonwords (McElroy & Slamecka, 1982; Nairne et al., 1985). Based on this finding, McElroy and Slamecka argued that the generation effect occurs because, relative to reading, generating an item produces greater activation of that item's representation in lexical memory. Thus, if a lexical representation of the generated item does not exist, as is the case for nonwords, no generation effect will be observed. Because support for this lexical activation hypothesis hinges largely on the failure to obtain a generation effect with nonwords (McElroy & Slamecka, 1982; Nairne et al., 1985), we sought an independent replication of this result in the present research. More importantly, we wanted to test an alternative interpretation of the word versus nonword difference in which lexicality is not viewed as a critical factor for obtaining the generation effect. This alternative interpretation is viable because of the procedural differences, which we describe in more detail below, that exist between the word conditions that have produced the generation effect and those nonword conditions that have not.

McElroy and Slamecka (1982, Experiment 1) presented subjects with both word and nonword pairs. For the word pairs, the stimulus and response items were related by

---

Portions of this research were reported at the annual meeting of the Midwestern Psychological Association held in Chicago in May, 1985. This research was supported in part by summer faculty grants from the Research Foundation of the State University of New York, to the first author, and from Purdue Research Foundation, to the second author. Requests for reprints should be sent to D. G. Payne, Department of Psychology, State University of New York at Binghamton, Binghamton, NY 13901.

an antonym rule; for the nonword pairs, they were related by a letter-transposition rule in which a subset of the letters from the stimulus could be rearranged to form the nonword response. A subsequent recognition memory test yielded a generation effect for the word responses but not for the nonword responses. In two additional experiments, McElroy and Slamecka (1982) assessed the reliability of the absence of a generation effect for nonwords and ruled out some plausible, but rather uninteresting, interpretations of the absence. One such interpretation is that the absence of a generation effect for nonwords depends on embedding the nonwords in a list containing word pairs. This possibility was ruled out by McElroy and Slamecka's (1982, Experiment 2) failure to find a generation effect when subjects studied only nonword pairs, a finding subsequently replicated by Nairne et al. (1985, Experiments 1 and 2). McElroy and Slamecka (1982, Experiment 3) also demonstrated that their failure to obtain a generation effect with nonwords was not due to some unusual aspect(s) of the letter-transposition rule. They did this by showing that a rhyme rule also fails to yield a generation effect for nonword pairs, another finding that has been recently replicated by Nairne et al. (1985). Since a rhyme-generation rule has been shown to yield a generation effect for words (e.g., Slamecka & Graf, 1978), McElroy and Slamecka (1982) argued that it was the nonlexicity of the response term and not the type of generation rule that caused the absence of a generation effect for their nonwords.

One limitation of McElroy and Slamecka's (1982) and Nairne et al.'s (1985) experiments is that these researchers confounded the lexicality of the stimulus and response items. Hence, the absence of a generation effect for nonwords could have been due to either the lexicality of the response item (as McElroy and Slamecka argued) or the lexicality of the stimulus item, or to both. This latter possibility provides the basis for an alternative hypothesis—hereafter called the redintegration hypothesis (see Horowitz & Manelis, 1972; Horowitz & Prytulak, 1969)—as to why McElroy and Slamecka found a generation effect for words but not for nonwords. According to this hypothesis, correct recognition can be based on either of two separate recognition processes (see Atkinson & Juola, 1973; Gillund & Shiffrin, 1984; Mandler, 1980). The first of these is based on recognizing the test item in isolation and depends on the test item's familiarity. The second, redintegrative process is a recall-like process in which subjects use the test item to retrieve contextual information, including other list items.

To account for McElroy and Slamecka's (1982) results, we assume that the generation effect is localized in the second, redintegrative process in which subjects use the generation rule to try to retrieve the stimulus term with which the response test item was studied (see Bartling & Thompson, 1977; Salzberg, 1976). Presumably, the probability of successful stimulus retrieval in the presence of the response test item is inversely proportional to the total number of possible stimulus items related to that response

item via the specified generation rule (see Nelson & McEvoy, 1979). Because the generation effect is localized in the redintegration process and the probability of retrieving the stimulus term during that process should be much higher for word-word pairs (related via a relatively specific antonym rule) than for nonword-nonword pairs (related by the much more general letter-transposition rule), the generation effect should be much greater for the word-word pairs. Thus, the redintegration hypothesis holds that it is the nature of the *stimulus* item and its relation to the response that determine the presence or absence of a generation effect. This is in contrast to the lexical activation hypothesis, which maintains that it is the lexical status of the *response* term that is critical.<sup>1</sup>

This redintegration hypothesis can also account for the results of an experiment McElroy and Slamecka (1982) mentioned briefly in their discussion section as being problematic for their lexical activation hypothesis. In that experiment, a generation effect did not occur for response words when a rhyme rule and a nonword stimulus (e.g., *gomp-romp*) were used. Since the response term was a word, the lexical activation hypothesis predicts that a generation effect should have occurred. However, as argued above, the redintegration hypothesis can account for this result if, as seems likely, the probability that subjects can retrieve the studied nonword stimulus term to the response test word will be low because there are many nonwords that rhyme with the response-term test word (see Nelson & McEvoy, 1979).

The redintegration hypothesis also predicts that the generation effect should extend to the stimulus as well as the response item, for it is in the redintegration of the stimulus-response unit that the generation effect occurs. Although Slamecka and Graf (1978, Experiment 3) failed to find a statistically significant generation effect for the recognition of stimulus terms from word-word pairs, a generation effect was numerically present. Furthermore, Johnson and Greenwald (1985) recently reported a significant generation effect for stimulus terms from word-word pairs. Finally, the redintegration hypothesis can also account for the generation effect in recall if one assumes that during recall subjects implicitly try to redintegrate the stimulus-response units.

We sought to achieve two primary goals with the present research. First, we wanted to replicate McElroy and Slamecka's (1982) word versus nonword generation effect difference with procedures in which (1) the same rule is used to generate the word and nonword responses and (2) the stimulus terms for the word and nonword responses are the same (and hence equated on their lexicality). Second, we wanted to test between the lexical activation and the redintegration interpretations of the word versus nonword difference in the generation effect. To achieve these ends, we varied orthogonally the lexicality (word vs. nonword) of the stimulus and response items. The same rule (rhyme) was used to generate both word and nonword responses, and the stimulus and response differed only by their initial letter (e.g., *chop-shop*). Thus,

unlike previous researchers who compared the generation effect for words and nonwords, we equated our word and nonword responses in terms of (1) the lexicality of the stimulus terms with which they were paired, (2) the generation rule that relates them to these stimulus terms, and (3) their properties on the (phonological) dimension being tapped by that generation rule.

To maximize our chances of supporting the redintegration hypothesis, Experiment 1 included two other design features. First, to make it more likely that subjects would be better able to retrieve an integrated unit when the stimulus item was a word rather than a nonword (as was arguably the case in McElroy and Slamecka's, 1982, experiments), we used stimulus-response pairs that rhymed with very few other English words. Second, we manipulated the lexicality of the stimulus and response terms between subjects, such that all stimulus terms and all response terms a subject studied had the same lexical status. Thus, during redintegration, subjects could restrict their search for stimulus terms to rhyming words or nonwords. This should accentuate the difference in the probabilities that subjects can retrieve the word stimulus terms from a small potential set and the nonword stimulus terms from a large potential set (see Nelson & McEvoy, 1979).

Both the lexical activation and redintegration hypotheses predict a generation effect for word-word pairs but not for nonword-nonword pairs, which would replicate McElroy and Slamecka's (1982) results. However, because the lexical activation hypothesis predicts a generation effect only when the response term is a word, it predicts a generation effect for nonword-word pairs but not for word-nonword pairs. On the other hand, the redintegration hypothesis predicts the opposite, under the assumption that the generation effect is localized in a stimulus redintegration process, the success of which Experiment 1 was designed to make much more likely for word than for nonword stimulus terms.

## EXPERIMENT 1

### Method

**Subjects and Design.** Subjects were 168 Purdue University introductory psychology students participating in partial fulfillment of a course requirement for research participation or library research. A  $2 \times 2 \times 2$  mixed factorial design was used with stimulus type (word vs. nonword) and response type (word vs. nonword) as between-subjects factors and task (generate vs. read) as a within-subjects factor.

**Materials.** A pool of 192 rhyming word pairs was selected using two criteria. First, the words in each pair differed only by the initial letter (e.g., *chop-shop*). Second, for each pair of words selected, there were very few other English words that could be formed by changing the initial letter. These 192 word pairs were then assigned randomly to three lists (A, B, and C) of 64 pairs each. The words in these three lists did not differ in the number of additional English words that could be formed from them by changing the initial letters ( $F < 1$ ), with the means ranging from 1.00 to 1.35.

These three base lists were used to construct two alternate forms (I and II) of each study list. The order in which pairs were studied was held constant across the two forms; however, the assignment

of items to task (generate vs. read) differed. Half of the items from List A were assigned randomly to the generate task, and the remaining items were assigned to the read task (Form I). This assignment was reversed for Form II, and these same two task orders were then used to create two alternate forms of Lists B and C. These six word-word lists (3 lists  $\times$  2 forms) were then used to construct lists for the remaining three study conditions (i.e., word-nonword, nonword-word, and nonword-nonword) by changing the initial letter of the stimulus and/or response term of the pair so as to produce a pronounceable nonword (e.g., *chop-thop*, *phop-shop*, *phop-thop*). Seven subjects were assigned to each of the 24 study list conditions (2 stimulus item types  $\times$  2 response item types  $\times$  3 lists  $\times$  2 forms).

The 64 item pairs from each list were typed in lowercase letters on 3-in. (7.62-cm)  $\times$  5-in. (12.7-cm) index cards. For the generate task, the stimulus term was followed by the initial letter of the target response item and blank lines corresponding to the letters missing from the response term (e.g., *chop-s\_ \_*). For the read task, both the stimulus and response appeared intact.

Two alternate forms of a three-alternative forced-choice (3AFC) recognition test were constructed, one for the word response terms and one for the corresponding nonword response terms. Each test contained 64 lines with three items per line, one from each list (A, B, C); the order of A, B, and C items within a line was varied randomly across the 64 triads. Each triad of items contained response items from the same serial input position within their respective study lists. Thus, each triad contained all nonrhyming words or nonwords. However, for both words and nonwords, in a few triads, a nonstudied item sometimes rhymed with another item from the study list.

**Procedure.** Individually tested subjects were informed of the lexical status (word vs. nonword) of the items in the study list and that the list contained two different types of study pairs (read and generate). They were told that read pairs contained a word (nonword) followed by a rhyming word (nonword) and that they should read both items aloud. Instructions indicated that the generate pairs contained a word/nonword followed by a letter and blank lines, and that the subject's task was to read the first item aloud and then say aloud a word/nonword that began with that letter and rhymed with the first item. Subjects were also told that they would subsequently receive a recognition test for the second item from each studied pair.

Subjects self-presented item pairs at a 5-sec rate and were instructed to turn to the next card in the deck each time they heard the prerecorded auditory signal ("NEXT"). Subjects were given a short practice list containing pairs appropriate to the condition to which they had been assigned (e.g., word-word). This practice list was followed by the critical list presentation. Subjects were then allowed as much time as needed to complete the 3AFC recognition test.

### Results and Discussion

For this and each of the following experiments, items for which errors occurred during input were not scored in the recognition test. (Replicating previous studies, errors during the input phase were virtually nonexistent—e.g., in Experiment 1, range = .04–.78 items per condition out of a possible 32.) The top portion of Table 1 displays the mean proportion correct recognition scores for the generate and read tasks for each of the four study conditions. These data were analyzed using a 2 (stimulus type: word vs. nonword)  $\times$  2 (response type: word vs. nonword)  $\times$  2 (task: read vs. generate) mixed factor analysis of variance with repeated measures on the last factor. (Unless otherwise noted, all effects called significant had two-tailed  $p$  values  $< .05$ .)

**Table 1**  
**Mean Proportion Correct Recognition as a Function of**  
**Encoding Condition and Item Type in Experiments 1-3**

| Item Type<br>Stimulus-Response | Encoding Condition |          | Difference |
|--------------------------------|--------------------|----------|------------|
|                                | Read               | Generate |            |
| Experiment 1                   |                    |          |            |
| Word-Word                      | .77                | .86      | .09*       |
| Nonword-Word                   | .74                | .82      | .08*       |
| Word-Nonword                   | .70                | .72      | .02        |
| Nonword-Nonword                | .67                | .68      | .01        |
| Experiment 2                   |                    |          |            |
| Word-Word                      | .74                | .88      | .14*       |
| Nonword-Word                   | .73                | .84      | .11*       |
| Word-Nonword                   | .69                | .74      | .05        |
| Nonword-Nonword                | .73                | .76      | .03        |
| Experiment 3                   |                    |          |            |
| Word-Word                      | .81                | .86      | .05†       |
| Nonword-Word                   | .79                | .90      | .11*       |
| Word-Nonword                   | .74                | .73      | -.01       |
| Nonword-Nonword                | .71                | .74      | .03        |

\* $p < .05$ , two-tailed. † $p < .06$ , two-tailed.

There were significant main effects of both stimulus type [ $F(1,164) = 5.92$ ,  $MSe = .018$ ] and response type [ $F(1,164) = 46.89$ ,  $MSe = .018$ ]. Thus, in Experiment 1, regardless of whether the responses were generated or read, word stimuli and word responses produced better performance than did nonword stimuli and nonword responses. This conclusion regarding the effect of the lexical status of the stimulus items should be viewed with caution, however, as we did not obtain a significant main effect of stimulus type when we manipulated this factor as a within-subjects factor (Experiments 2 and 3). Although the main effect of task was significant [ $F(1,164) = 31.07$ ,  $MSe = .007$ ], not all conditions produced a generation effect. Specifically, the generation effect was modulated by the lexicality of the response [ $F(1,164) = 14.04$ ,  $MSe = .007$ , for the task  $\times$  response type interaction] but not by the lexicality of the stimulus ( $F < 1$ , for the task  $\times$  stimulus type interaction). Planned comparisons showed the generation effect to be significant for the word-word [ $t(41) = 5.66$ ] and nonword-word [ $t(41) = 4.09$ ] pairs but not for the word-nonword [ $t(41) = 1.20$ ] or nonword-nonword [ $t(41) = .54$ ] pairs.

The findings that a generation effect occurred for nonword-word pairs but not for word-nonword pairs are contrary to our redintegration hypothesis, despite the fact that Experiment 1 was designed to maximize our chances of supporting it. However, the results agree completely with McElroy and Slamecka's (1982) lexical-activation hypothesis in that a generation effect occurred only for word response terms and regardless of whether the stimulus term was a word or a nonword.

## EXPERIMENT 2

Though congruent with the lexical activation hypothesis, our finding of a generation effect for nonword-word pairs runs counter to McElroy and Slamecka's (1982) in-

formally reported failure to find a generation effect for such pairs. Thus, the first goal of Experiment 2 was to replicate the generation effect we obtained with nonword-word pairs in Experiment 1. A second goal was to provide an additional test of the lexical activation hypothesis. Experiment 1 had the nice feature of directly comparing generation effects for (phonologically) "matched" word and nonword responses that were related to exactly the same stimulus terms via the same (i.e., rhyme) rule. However, because the response term's lexicality was manipulated between subjects, it may still be that a generation effect will be obtained for nonword responses generated from a rhyme rule if some of the response terms in the list are words. That is, the presence of word responses in the list might encourage subjects to first consult lexical memory in their attempt to generate the rhyming response. Perhaps the generation effect will occur whenever this initial consultation with lexical memory occurs, whether or not subjects successfully locate the item in lexical memory.<sup>2</sup> Experiment 2 examined this possibility by replicating Experiment 1 with a within-subjects design. If the generation effect depends on the particular generated response term's having a representation in lexical memory, as McElroy and Slamecka have implied, Experiment 2 will once again show a generation effect for word responses but not for nonword responses. If, on the other hand, the generation effect depends on subjects' consulting of lexical memory while generating their responses and the presence of some word responses in the list induces this consultation, then Experiment 2 will show a generation effect for nonword responses as well as word responses.<sup>3</sup>

## Method

Except where noted, the design and procedures of Experiment 2 were identical to those of Experiment 1. Subjects were 48 introductory psychology students at the State University of New York at Binghamton and all factors of the  $2 \times 2 \times 2$  design were now manipulated within subjects. The study items from Experiment 1 were assigned to new lists in Experiment 2 such that each study list contained 16 pairs of each of the four different pair types (e.g., word-nonword). Eight pairs of each type were assigned to the generate condition and 8 to the read condition. The assignment of pairs to each of the eight conditions (4 pair types  $\times$  2 tasks) was counterbalanced across lists, with 2 subjects being assigned to study a particular version of the 24 different study lists (eight conditions counterbalanced for each of 3 study lists). Subjects were informed that the stimulus and response terms could be words or nonwords. The 3AFC recognition tests from Experiment 1 were changed so that each test was now composed of half word triads and half nonword triads corresponding to the studied items from the same serial position within their respective lists.

## Results and Discussion

As can be seen in the middle portion of Table 1, the results of Experiment 2 are in close agreement with those of Experiment 1. There was a significant main effect of response type [ $F(1,47) = 23.6$ ,  $MSe = .016$ ], indicating that recognition was better for word than for nonword responses. The generation effect was also significant [ $F(1,47) = 47.8$ ,  $MSe = .009$ ]. More importantly,

however, there was once again a significant task  $\times$  response type interaction [ $F(1,47) = 10.3$ ,  $MSe = .017$ ]. Replicating Experiment 1, a generation effect was obtained when the target response item was a word [ $t_s(47) = 3.39$  and  $8.41$  for word-word and nonword-word pairs, respectively] but not when it was a nonword [ $t_s(47) = 1.74$  and  $.77$  for word-nonword and nonword-nonword pairs, respectively].

### EXPERIMENT 3

Because only 50% of the response terms in the study lists of Experiment 2 were words, subjects may not have been induced to consult lexical memory as they had been for the word-word and nonword-word lists of Experiment 1, in which 100% of the response terms were words. To examine generation effects for word and nonword response terms studied by the same subjects under identical conditions designed to maximize the probability that subjects would consult lexical memory, in Experiment 3, we replicated Experiment 2 except that we required subjects to make a lexical decision to each response term. Perhaps by forcing subjects to consult lexical memory via this procedure, we will be able to observe a generation effect for nonwords.

#### Method

Experiment 3 differed from Experiment 2 in two regards: A new group of 48 Binghamton students was recruited, and after subjects said aloud each response term, they then said aloud whether it was a word or a nonword.

#### Results and Discussion

The overall pattern of results in Experiment 3 (see Table 1) was very similar to that obtained in Experiments 1 and 2. Word responses were better recognized than nonword responses [ $F(1,47) = 51.08$ ,  $MSe = .022$ ], and a generation effect occurred [ $F(1,47) = 9.58$ ,  $MSe = .020$ ]. However, as before, the generation effect interacted with the effects of response type [ $F(1,47) = 6.09$ ,  $MSe = .022$ ] such that the generation effect occurred for word responses but not for nonword responses. The post hoc comparisons for the generation effects for the four pair types yielded the same pattern of statistical significance as before, except that the generation effect for the word-word pairs was only marginally significant [ $t(47) = 1.87$ ,  $p < .06$ , two-tailed]. This was due to the fact that performance in the read condition, relative to that observed in Experiment 2, was enhanced somewhat by the requirement that subjects make lexical decisions to the target. At first glance, this result seems to indicate that the lexical decision task served to make the read condition functionally more similar to the generate condition. This seems unlikely, however, since there was no similar relative enhancement of performance in the read condition for the nonword-word pairs. Given the consistent pattern of data in Experiments 1-3, it seems safe to conclude that the somewhat diminished generation effect observed with the word-word pairs in Experiment 3 is due to sampling er-

ror. The important point remains, however, that forcing subjects to consult lexical memory by requiring them to make a lexical decision to each response item was not sufficient to produce a generation effect for nonword responses.

### GENERAL DISCUSSION

The main result obtained in all three of our experiments was that a significant generation effect occurred whenever the response item was a word but not when it was a nonword. These results are completely consistent with McElroy and Slamecka's (1982) lexical activation hypothesis, which maintains that a generation effect will only be obtained when the generated target item is represented in the subjective lexicon. Furthermore, the data from our experiments offer no support for a reintegration hypothesis, which holds that the nature of the stimulus item and its relation to the response item play a role in determining the presence or absence of a generation effect by affecting the probability that the stimulus term will be reintegrated during the recognition test for the response item. That is, in no case was the generation effect influenced by the lexicality of the stimulus item. Finally, the results of Experiments 2 and 3 indicate that inducing the subject to consult lexical memory is not sufficient for producing a generation effect for nonwords. This finding replicates and extends the results of Nairne et al. (1985, Experiment 1), who found no generation effect with nonwords even when subjects were told that these items were actually low-frequency English words.

The present experiments clearly establish that a generation effect occurs for words but not for nonwords, even when the words and nonwords are phonologically "matched" and generated using the same rule (Experiments 1-3). Furthermore, these same results are obtained under conditions likely to induce general lexical-consultation strategies for both the word and nonword items (Experiments 2-3). Nevertheless, it remains unclear which of the many variables confounded with response lexicality underlie this word/nonword difference in the generation effect. Indeed, the results of two recent studies indicate that lexical status of the target item per se is not the only factor influencing the generation effect. Specifically, Gardiner and Hampton (1985) reported a generation effect when the target response item was a familiar compound noun (e.g., *cheese cake*) but not when it was an unfamiliar "compound noun" (e.g., *cheese ketchup*), whereas Nairne et al. (1985, Experiment 3) found a significant generation effect for high- and medium-frequency words but not for low-frequency words or nonwords. These findings, along with the word/nonword difference in the generation effect, suggest that although representation in the subjective lexicon may be a necessary condition for producing the generation effect, it is not a sufficient one.

Gardiner and Hampton (1985) argued that the critical factor mediating their generation effects was that the target response form an integrated functional unit in memory.

On the other hand, under the reasonable assumption that high- and medium-frequency words have several meanings whereas low-frequency words and nonwords have very few (or no) meanings, Nairne et al. (1985) concluded that the critical factor mediating their generation effects was that the response term have some minimal number of different meanings. However, Nairne et al.'s results are equally congruent with Gardiner and Hampton's (1985) unitization hypothesis, assuming that high- and medium-frequency words are more unitized than are low-frequency words or nonwords (see Healy, 1976). Clearly, additional analytical research must be conducted to establish the validity of these (or other) interpretations of the role that response lexicality plays in the generation effect. Nevertheless, the present results just as clearly demonstrate that the word/nonword difference in the generation effect can be obtained under conditions in which one (1) uses identical stimulus items (hence equating them on their lexicality), (2) uses the same generation rule (i.e., rhyme) for the word and nonword responses, and (3) matches the word and nonword responses on the dimension (i.e., phonological) on which the word and nonword responses are related to the stimulus items via that rule.

#### REFERENCES

- ANDERSON, R. C., GOLDBERG, S. R., & HIDDE, J. L. (1971). Meaningful processing of sentences. *Journal of Educational Psychology*, *62*, 395-399.
- ATKINSON, R. C., & JUOLA, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press.
- BARTLING, C. A., & THOMPSON, C. P. (1977). Encoding specificity: Retrieval asymmetry in the recognition failure paradigm. *Journal of Experimental Psychology: Human Learning & Memory*, *3*, 690-700.
- 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.
- GARDINER, J. M., & ROWLEY, J. M. C. (1984). A generation effect with numbers rather than words. *Memory & Cognition*, *12*, 443-445.
- GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1-67.
- 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 intraword 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.
- HEALY, A. F. (1976). Detection errors on the word *the*: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception & Performance*, *2*, 235-242.
- HOROWITZ, L. M., & MANELIS, L. (1972). Toward a theory of redintegrative memory: Adjective-noun phrases. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 193-224). New York: Academic Press.
- HOROWITZ, L. M., & PRYTULAK, L. S. (1969). Redintegrative memory. *Psychological Review*, *76*, 519-531.
- JACOBY, L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning & Verbal Behavior*, *17*, 649-667.
- JACOBY, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, *22*, 485-508.
- JOHNSON, M. M. S., & GREENWALD, A. G. (1985, May). Does the generation effect apply to stimuli as well as responses? Paper presented at the 57th annual meeting of the Midwestern Psychological Association, Chicago.
- KANE, J., & ANDERSON, R. C. (1978). Depth of processing and interference effects in the learning and remembering of sentences. *Journal of Educational Psychology*, *70*, 626-635.
- MANDLER, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252-271.
- 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*, 243-259.
- MCFARLAND, C. E., FREY, T. J., & RHODES, D. D. (1980). Retrieval of internally versus externally generated words in episodic memory. *Journal of Verbal Learning & Verbal Behavior*, *19*, 210-225.
- NAIRNE, J. S., PUSEN, C., & WIDNER, R. L. (1985). Representation in the mental lexicon: Implications for theories of the generation effect. *Memory & Cognition*, *13*, 183-191.
- NELSON, D. L., & MCEVOY, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning & Memory*, *5*, 292-314.
- SALZBERG, P. M. (1976). On the generality of encoding specificity. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 586-596.
- SLAMECKA, N. J., & FEVREISKI, J. (1983). The generation effect when generation fails. *Journal of Verbal Learning & Verbal Behavior*, *22*, 153-163.
- SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 592-604.
- STEIN, B. S., & BRANSFORD, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning & Verbal Behavior*, *18*, 769-777.
- WINNICK, W. A., & DANIEL, S. A. (1970). Two kinds of response priming in tachistoscopic recognition. *Journal of Experimental Psychology*, *84*, 74-81.

#### NOTES

1. Our redintegration hypothesis is superficially similar to Donaldson and Bass's (1980) solution-checking explanation of the generation effect. According to them, because there is uncertainty as to the correctness of a response when subjects generate the response from incomplete information, subjects scan back to the stimulus term to check if the response they generated is indeed correct. A stronger memory trace for a generated response term is produced by this redintegrative checking operation, which is not performed in the read condition because the correct response is fully available in the completely spelled-out word. Although we too assume a redintegrative check involving the stimulus term, we assume it occurs at retrieval, rather than at storage, as Donaldson and Bass assume. Also, unlike our explanation, Donaldson and Bass's explanation has trouble accounting for the lack of a generation effect for nonwords, since it would seem that subjects would need to check the correctness of a generated nonword at least as much as they would need to check the correctness of a generated word.
2. The results of two published studies are relevant to the idea that the generation effect depends not upon the response item being represented in lexical memory but rather upon the subject's merely at-

tempting to locate the response item in lexical memory. This lexical consultation hypothesis receives indirect support from Slamecka and Fevreski's (1983) finding that if a subject attempts unsuccessfully to generate a target word from a stimulus word using an antonym rule, a generation effect is still obtained in recognition whether or not that target word is provided by the experimenter following the unsuccessful generation attempt. However, this lexical consultation hypothesis was not supported by results from Nairne et al.'s (1985) Experiment 1. Although their subjects received all nonword-nonword pairs, related by a rhyme rule, Nairne et al. attempted to encourage them to consult lexical memory by telling them that the items they would be seeing were all very low-frequency words. A generation effect was not obtained for nonwords under these conditions. However, we believe the inclusion of easily recognized words in our experiment should, relative to Nairne et al.'s instructional manipulation, provide a much stronger inducement for subjects to consult lexical memory.

3. Although McElroy and Slamecka (1982) have already shown that a generation effect does not occur for nonword responses when word responses are present in the same study list, this result was obtained when word stimuli and an antonym rule were used for word responses and when nonword stimuli and a letter transposition rule were used for nonword responses. Thus, subjects may have attempted to retrieve their responses from lexical memory when a word stimulus occurred but not when a nonword stimulus occurred. Such a strategy cannot be employed in the present experiment, because the lexicality of the response is not predictable from the lexicality of the stimulus term.

(Manuscript received July 30, 1985;  
revision accepted for publication January 6, 1986.)