

A dissociative word-frequency × levels-of-processing interaction in episodic recognition and lexical decision tasks

JANET M. DUCHEK

Washington University, St. Louis, Missouri

and

JAMES H. NEELY

State University of New York at Albany, Albany, New York

The effects of levels-of-processing and word frequency were directly compared in three different memory tests. In the episodic recognition test, the subjects decided whether or not a word or a pronounceable nonword had been previously studied. In the two lexical decision tests with either pronounceable or unpronounceable nonwords as distractors, the subjects decided whether a test item was a word or a nonword. There were four main results: (1) in all three tests, reaction times (RTs) in response to studied words were faster if they had received semantic rather than rhyme processing during study; (2) in the episodic recognition test, RTs were faster for low- than for high-frequency words; in both lexical decision tests, RTs were faster for high- than for low-frequency words, though less so when the nonword distractors were unpronounceable; (3) prior study facilitated lexical decisions more in response to low- than to high-frequency words, thereby attenuating the word-frequency effect, but more so when the nonword distractors were pronounceable; (4) in the lexical decision test with pronounceable nonword distractors, relative to prior rhyme processing, prior semantic processing facilitated performance more for high- than for low-frequency words, whereas the opposite was the case in the episodic recognition test. Discussion focused on the relationship of these results to current views of the mechanisms by which (1) word frequency and depth of processing affect performance in implicit and explicit memory tests, and (2) repetition priming attenuates word-frequency effects for lexical decisions.

Recently, there has been a burgeoning interest in comparing performance in explicit and implicit memory tests (e.g., Graf & Schacter, 1985; Richardson-Klavehn & Bjork, 1988; Schacter, 1987; Tulving, 1985). In explicit memory tests, subjects indicate whether they remember having studied an event at a particular time and/or place. For example, when explicit recognition memory for words is assessed, subjects are to respond "yes" to a test word if they remember it as having been studied in the ex-

perimental context. Memory is considered good to the degree that subjects respond "yes" to studied test words more often than to nonstudied test words. In the analogous implicit memory test, subjects respond to a test item that corresponds to a word that has or has not been previously studied in the experimental context, without being required to recollect its prior study status. For example, in an implicit memory test, subjects might identify a briefly presented, masked test word, or respond with a word that completes a word fragment with missing letters. Implicit memory is considered better to the degree that subjects perform better on test items corresponding to previously studied words, relative to test items corresponding to nonstudied words—an effect typically referred to as priming.

A major focus of interest has been whether dissociation effects occur for implicit and explicit memory tests. A dissociation effect occurs whenever a variable that enhances memory in an explicit memory test has either no effect or a deleterious effect on memory in a corresponding implicit memory test, and vice versa. In the present research, we consider two such dissociation effects, levels-of-processing effects and word-frequency effects, which we first briefly review.

Both authors contributed equally. This research was conducted by Janet M. Duchek in partial fulfillment of the requirements for an M.S. degree at the University of South Carolina, under the direction of James H. Neely. Preparation of the manuscript was supported in part by an XL grant from Purdue Research Foundation and by NICHHD Grant R01 HD15054, both to James H. Neely. We thank David Clement, who served on the M. S. committee, and especially Randall Engle, who served as the chairman in the second author's absence. We also thank Dave Balota, Larry Jacoby, Gordon Logan, and Henry L. Roediger III for their useful comments. This research was reported at the annual meeting of the Southeastern Psychological Association, held in Atlanta in March, 1979. Requests for reprints should be sent to either Janet M. Duchek, Department of Psychology, Washington University, St. Louis, MO 63130, or James H. Neely, Department of Psychology, State University of New York at Albany, Albany, NY 12222.

LEVELS-OF-PROCESSING EFFECTS IN IMPLICIT AND EXPLICIT MEMORY TESTS

Although explicit memory of words is much better when they have previously been processed to a "deep" semantic level rather than a "shallow" graphological or phonological level (see Craik & Lockhart, 1972; Craik & Tulving, 1975; Graf & Mandler, 1984; Graf, Mandler, & Haden, 1982; Graf, Squire, & Mandler, 1984; Jacoby & Dallas, 1981), levels-of-processing effects are less clear-cut in implicit memory tests. For example, in a perceptual identification test with briefly presented masked words, Jacoby & Dallas (1981) have shown that the priming produced by a test word's prior unmasked presentation remains the same, regardless of whether that prior presentation has been processed to a deep or to a shallow level. However, in implicit memory tests that involve word-stem completion, in which the initial letters of a word serve as a retrieval cue, deeper prior processing yields slightly greater priming—though within individual experiments, this levels-of-processing effect sometimes reaches statistical significance (Graf et al., 1984; Squire, Shimamura, & Graf, 1987) and sometimes does not (Graf & Mandler, 1984; Graf et al. 1982; Roediger, Weldon, & Stadler, 1987).¹

When a lexical (word/nonword) decision task is used as the implicit memory test, the results are even less clear. Kirsner, Milech, and Standen (1983) manipulated levels of processing between subjects during a first test phase. In a second test phase, the subjects performed a perceptual identification task or a lexical decision task (LDT) on masked test items presented for 40 msec, or an LDT on nonmasked test items presented for 2 sec. In a third test phase, all the subjects performed an explicit recognition test on nonmasked test items presented for 3 sec. Although the explicit recognition test yielded the standard levels-of-processing effect for both speed and accuracy, the implicit memory tests yielded mixed results. Specifically, the conclusion as to whether or not a levels-of-processing effect was obtained depended on whether one only considered performance on the studied items, or measured priming from prior study relative to performance on nonstudied items. This was so because of large differences in performance on nonstudied items (for which the levels-of-processing manipulation was a pseudo-manipulation) for the groups of subjects who performed deep as opposed to shallow processing during the study phase. Hayman (1983) avoided this problem of different performance levels on nonstudied items by manipulating levels of processing within subjects. He found faster reaction times (RTs) in response to words previously processed to a deep rather than a shallow level, but only when the instructions emphasized accuracy rather than speed. The research reported here will provide more data on the issue of whether levels-of-processing effects can occur for implicit memory when it is assessed in an LDT.

The present research also provides a test of data-driven versus conceptually driven processing as an account of

why levels-of-processing effects are smaller for implicit than for explicit memory measures—an account that has been proffered in various guises by several researchers (e.g., Blaxton, 1985; Graf & Mandler, 1984; Jacoby, 1983; Jacoby & Dallas, 1981; Roediger & Blaxton, 1987; Roediger et al., 1987). By this account, implicit memory tests typically invoke data-driven processing operations that tap an item's perceptual fluency or familiarity, which is little affected by prior processing depth. Explicit memory tests, on the other hand, typically invoke conceptually driven processing operations that tap information about an item's elaborative associations with other studied items, the study context, and other related items in semantic memory, and these elaborative associations are affected by prior processing depth. The degree to which subjects rely on an item's familiarity as opposed to its elaborative associations depends on the particular testing conditions (e.g., see Johnston, Dark, & Jacoby, 1985; Morris, Bransford, & Franks, 1977; Roediger & Blaxton, 1987). The exact manner in which the present experiment provides a test of data-driven versus conceptually driven processing as an account of dissociative levels-of-processing effects in implicit and explicit memory tests will be given shortly.

WORD-FREQUENCY EFFECTS IN EXPLICIT AND IMPLICIT MEMORY TESTS

In explicit recognition memory tests, performance is better for low- than for high-frequency words (e.g., Balota & Neely, 1980; Glanzer & Adams, 1985; Glanzer & Bowles, 1976; Gregg, 1976; Jacoby & Dallas, 1981; Shepard, 1967); however, in implicit recognition memory tests that involve either lexical decisions to unmasked words or the identification of masked words, performance is better for previously presented high-frequency words than for previously presented low-frequency words (e.g., Forster & Davis, 1984; Jacoby & Dallas, 1981; Kirsner et al., 1983; Scarborough, Cortese, & Scarborough, 1977). But because the *priming* produced by a word's prior unmasked presentation is greater for low- than for high-frequency words (see the immediately preceding citations), Jacoby and Dallas (1981) have argued that implicit and explicit memory tests both show that memory for an item's prior presentation is better for low- than for high-frequency words.

Forster and Davis (1984) have extended Jacoby and Dallas's (1981) argument to account for the word-frequency attenuation effect in the LDT, which is the observation that the superiority in lexical decision performance for high-frequency words relative to low-frequency words is attenuated when these high- and low-frequency words have been previously presented. Since it is well-known that explicit episodic recognition memory is better for low- than for high-frequency words, Forster and Davis claim that the word-frequency attenuation effect—the greater benefit from repetition priming for low-frequency words than for high-frequency words—is be-

ing mediated by differences in explicit episodic memory for the low- and high-frequency words' first presentations. (We will discuss the empirical evidence for this claim later, when we describe how the present experiment directly tests it.)

EXPERIMENT

The present experiment examines whether levels-of-processing effects and word-frequency effects will be dissociated in explicit and implicit memory tests in which a previously studied test word is presented exactly as it was studied—that is, unmasked, with no letters missing. One important feature of the present experiment is that we administered our explicit and implicit memory tests after identical study phases, under testing conditions that were identical apart from the test instructions for making responses. As Neely and Payne (1983) and Neely (1989) have noted, this is important: when the two types of tests differ in variables other than instructions, it is unclear whether a variable's dissociated effects are due to its having different effects on the two types of memory being tested or to its effects' interacting with the effects of variables not controlled across the two types of test. Although Graf and Mandler (1984), Graf et al. (1982), Graf et al. (1984), and Roediger et al. (1987) have directly compared levels-of-processing effects in explicit and implicit (word-stem completion) cued *recall* tests equated on all variables other than the instructions, to our knowledge no one has used such well-controlled conditions to compare directly levels-of-processing effects or word-frequency effects in explicit and implicit *recognition* memory tests.

The present experiment consisted of two phases. In Phase 1, all the subjects responded to either a semantic or a rhyme question about each *study* item. In Phase 2, different groups of subjects performed either an LDT as the implicit memory test or an episodic recognition test as the explicit memory test. Subjects given the LDT with pronounceable nonword distractors (the LDT-PNW test) were told to press a "yes" key for a word and a "no" key for a nonword, regardless of whether or not it had appeared in the study phase; subjects given the episodic recognition (EPIS RGN) test were told to press a "yes" key if an item had appeared in the study phase and a "no" key if it had not, regardless of whether it was a word or a nonword.

In Phase 1, the semantic and rhyme questions were followed by an equal number of high- and low-frequency words and pronounceable nonwords derived from high- and low-frequency words. It was necessary to include pronounceable nonwords in both Phases 1 and 2 so that the LDT-PNW and EPIS RGN tests would (1) contain exactly the same test items, and (2) be administered after identical study phases. Specifically, it was necessary to include pronounceable nonwords in Phase 1 for LDT subjects; if only words had been included, the subjects could have used explicit memory to aid in their lexical decisions—namely, by responding "yes" ("word") to

any item they recognized as having occurred in Phase 1—thereby making the LDT an explicit rather than an implicit memory test (see Durgunoğlu & Neely, 1987; Neely & Durgunoğlu, 1985). We equated the number of "yes" ("no") responses in the two types of tests by letting the LDT-PNW and EPIS RGN tests contain equal numbers of studied and nonstudied words and nonwords. Finally, a third group of subjects was given an LDT with unpronounceable nonword distractors (the LDT-UNW test). Subjects receiving the LDT-UNW test were treated identically to those receiving the LDT-PNW test, apart from the fact that the nonwords that appeared in their test list were created by randomly permuting the letters that formed the pronounceable nonwords in the LDT-PNW test.

By testing the three groups under these conditions in the present experiment, we were able to address the three following questions concerning levels-of-processing effects and word-frequency effects in implicit and explicit memory tests.

(1) *Will levels-of-processing effects be similar or dissimilar in an explicit recognition memory test and an implicit LDT administered under identical conditions?* The two studies that provide data potentially relevant to this question (Hayman, 1983, and Kirsner et al., 1983) do not provide an unambiguous answer to it. Specifically, the results of the Kirsner et al. (1983) experiments were ambiguous because (1) their between-subjects levels-of-processing manipulation resulted in very different baseline performance levels on nonstudied words, for which the levels-of-processing manipulation was a pseudo-manipulation, and (2) their EPIS RGN and LDT tests differed in when they were given (the LDT was always given before EPIS RGN), and in whether or not they included nonwords. The Hayman (1983) experiments do not provide an unambiguous answer to this question (indeed, they were not intended to do so), because they did not include an EPIS RGN test. Because we manipulated levels-of-processing within subjects and included nonwords in both our explicit EPIS RGN and implicit LDT-PNW tests, which were administered at identical times after the study phase, we were able to avoid these problems.

(2) *If a levels-of-processing effect is obtained in the LDT, will its magnitude depend on the different types of processing induced by pronounceable and unpronounceable nonword distractors?* There is considerable evidence that in making their *word* lexical decisions, the subjects who receive an LDT-UNW test are faster and more likely to rely on data-driven processing (i.e., on an item's familiarity or perceptual fluency), and less likely to rely on conceptually driven processing (i.e., on an item's retrievability and semantic properties), relative to subjects who receive an LDT-PNW test (see Balota & Chumbley, 1984; James, 1975; Shulman & Davison, 1977; Shulman, Hornak, & Sanders, 1978). The present experiment therefore permits a test of whether levels-of-processing effects differ in implicit and explicit memory tests because they

are data-driven versus conceptually driven tests, respectively. Specifically, if this account is correct, one would expect that if a levels-of-processing effect were obtained in our LDTs, it should be smaller in the LDT-UNW test, in which performance is based primarily on data-driven processing, than in the LDT-PNW test, which is more akin to an explicit memory test in which performance depends more on conceptually driven processing.

(3) *Is explicit episodic memory responsible for repetition's attenuating the word-frequency effect in the LDT?* The data from the present experiment can also potentially provide a direct test of Forster and Davis' (1984) claim that episodic memory mediates the word-frequency attenuation effect. As noted earlier, they based this claim in part on the finding that explicit episodic recognition memory for a prior presentation of a low-frequency word is superior to that for a prior presentation of a high-frequency word (see Balota & Neely, 1980; Jacoby and Dallas, 1981). However, Forster and Davis (1984) provided additional support for their analysis in a series of experiments that showed that a word-frequency attenuation effect does not occur for masked or nonattended unmasked prior presentations of the high- and low-frequency words, even though these prior presentations may yield a small repetition priming effect. Presumably, these masked or nonattended prior presentations failed to produce a word-frequency attenuation effect because they produced only very weak episodic memory traces (see, e.g., Balota, 1983; Wolford & Morrison, 1980).

There are, however, two potential problems with Forster and Davis' (1984) analysis. The first one is that explicit episodic memory may not be involved in producing repetition priming in an LDT, even when the first presentation of the repeated item is unmasked and attended. Evidence for this claim comes from experiments by Scarborough et al. (1977) and Moscovitch (1982), who found that as the lag between the *study* (the first) presentation of an item and the *test* (the second) presentation of the item increased, there was little or no decrement in the amount of repetition priming that the first presentation produced for lexical decisions to the second presentation. This was so despite the fact that explicit episodic recognition memory for the study presentation dropped off dramatically as the lag between it and the test presentation increased. Even more striking was Moscovitch's finding that repetition priming effects for lexical decisions were just as large in amnesic subjects with virtually no explicit episodic memories as in subjects with normal explicit episodic memories.

A second problem with Forster and Davis' (1984) analysis of the word-frequency attenuation effect for lexical decisions is that their experiments did not provide an independent measure of the subjects' explicit episodic memories for high- and low-frequency words tested under the same conditions as those that yielded the word-frequency attenuation effect in their LDT. Indeed, when Scarborough et al. (1977) examined word-frequency effects as a function of repetition, using exactly the same

materials in both a continuous EPIS RGN task (Experiment 5) and an LDT (Experiment 2) administered under identical testing conditions, they found a word-frequency attenuation effect in the LDT even though their low-frequency words did not yield better explicit memory than the high-frequency words in their EPIS RGN test. Taken at face value, Scarborough et al.'s data fail to confirm Forster and Davis' (1984) claims that (1) subjects use episodic memory when making lexical decisions with respect to previously studied items, and (2) repetition attenuates the word-frequency effect in the LDT, because low-frequency words' prior presentations are more strongly represented in episodic memory than high-frequency words' prior presentations are. However, the subjects in Scarborough et al.'s continuous EPIS RGN task made episodic recognition judgments in response to the high- and low-frequency words' first presentations, whereas the subjects in their LDT made lexical decisions in response to them. If Scarborough et al. had required their subjects to encode the test items' first presentations in the same fashion prior to their EPIS RGN test and their LDT, their results might have favored Forster and Davis' analysis.

The present experiment's results can provide a direct test of Forster and Davis' (1984) explanation of the word-frequency attenuation effect if (1) the levels-of-processing manipulation differentially affects memory performance on high- and low-frequency words in the EPIS RGN test, and (2) a levels-of-processing effect occurs in the LDT-PNW test. Since these two tests are administered under identical conditions, Forster and Davis' analysis predicts that to the degree that the recognition of low-frequency words is greater than the recognition of high-frequency words in the EPIS RGN test, following rhyme (or semantic) processing, the word-frequency attenuation effect in the LDT-PNW test should also be greater following rhyme (or semantic) processing.

Method

Materials and list construction. The Kučera and Francis (1967) norms were used to select 112 high-frequency (> 50 per million) and 112 low-frequency (< 5 per million) words. Pronounceable nonwords were then made from 56 of the high-frequency words and 56 of the low-frequency words, by changing single letters in randomly determined positions in them. A rhyme question and a semantic question were then constructed for each of the remaining 56 high-frequency and 56 low-frequency words, such that the correct response to them would be "yes." The semantic question followed the form of "The *blank* chased the cat" (for the study item *dog*, for example), and the rhyme question had the form of "Does it rhyme with moon?" (for the study item *spoon*). A rhyme question for which the correct answer was "yes" was also constructed for each nonword; the semantic question constructed for each nonword did not have as its correct answer the word from which that nonword had been derived. (Obviously, the correct answer was always "no" for the semantic nonword questions.)

These materials were used to construct two nonoverlapping base lists for presentation in Phase 1. Each base list consisted of 28 high-frequency words, 28 low-frequency words, 28 high-frequency nonwords, and 28 low-frequency nonwords. A randomized presentation order was created for each base list; this order was used for all subjects who saw that list in Phase 1. Four versions of each of

these base lists were made by changing the question that preceded each item on the base list, such that across these four versions each word was preceded once by each of the four question types (i.e., semantic "yes," semantic "no," rhyme "yes," and rhyme "no"). For the words, the semantic and rhyme "no" questions were randomly selected "yes" questions of the same type associated with words assigned to the opposite processing condition. For example, if *dog* were serving in the rhyme "yes" condition, its semantic "yes" question (i.e., "The *blank* chased the cat") served as a semantic "no" question for some other word (e.g., *cone*), and if *bunk* were serving in the semantic "yes" condition its rhyme "yes" question (i.e., "Does it rhyme with skunk?") served as the rhyme "no" question for some other word (e.g., *sign*). As for the nonwords, across the four versions of each base list each nonword was preceded once by a rhyme "yes" and rhyme "no" question, and twice by a semantic "no" question. Thus, within each version of each Phase 1 base list, the subjects heard a particular question and saw a particular word only once. There were 7 instances each of the four question types for high-frequency words and for low-frequency words, and 7 instances each of the rhyme "yes" and rhyme "no" questions and 14 instances each of the semantic "no" questions for the high-frequency and low-frequency nonwords. Each version of each Phase 1 base list began with 3 primacy buffer trials and ended with 3 recency buffer trials. (The six different items that served as buffer trials were the same for all Phase 1 lists.) Thus, across the eight different Phase 1 lists (2 base lists \times 4 versions), each word and nonword occurred equally often as a studied and nonstudied item, and as a studied item it occurred equally often in each of its four processing conditions (with the exception of the nonwords, which occurred twice as often in the semantic "no" processing condition as in the rhyme "yes" and rhyme "no" conditions).

A single randomly ordered presentation of the 224 words and nonwords from both base lists served as the EPIS RGN and LDT-PNW Phase 2 tests. The Phase 2 LDT-UNW test was derived from the other two tests by replacing each nonword with a random permutation of its letters that created an orthographically illegal, unpronounceable nonword. Thus, as noted above, for the 112 items that had been studied in Phase 1, there were 7 instances of each of the semantic "yes," rhyme "yes," and rhyme "no" processing conditions for each item type (high- and low-frequency word and nonword), and 14 instances each of the semantic "no" processing condition for the high- and low-frequency nonwords. (Of course, for the LDT-UNW test, the "studied" nonwords had not actually been studied in their unpronounceable test form.) For the 112 Phase 2 nonstudied test items, there were 28 instances of each of the four item types (high- and low-frequency words and nonwords). Thus, in Phase 2, for all three tests, 112 of the test items required a "yes" response and 112 test items required a "no" response. For the EPIS RGN test, "yes" was the correct response to the high- and low-frequency words and nonwords that had been studied in Phase 1, and "no" was the correct response to the nonstudied high- and low-frequency words and nonwords. For the LDT-PNW and LDT-UNW tests, "yes" was the correct response to the high- and low-frequency words that had or had not been studied in Phase 1, and "no" was the correct response to the high- and low-frequency nonwords that had or had not been studied in Phase 1.

Procedure. Prior to Phase 1, individually tested subjects were told they were being tested on how quickly they could answer questions about words and nonwords. Oral instructions informed them about the nature of the Phase 1 questions and how they were to respond to them. (No mention of Phase 2 was made.) The subjects were told to look at a fixation point in the tachistoscope while the experimenter read the question. After hearing the question, the subjects pressed a telegraph key that simultaneously exposed the stimulus item at the fixation point for 3 sec and activated a response timer accurate to the nearest millisecond. The subjects responded to this item by using the same finger with which they initiated the exposure

of the test item to press either a "yes" or "no" response key; when depressed, the key stopped the response timer. The left and right positioning of the "yes" and "no" keys relative to the telegraph key that initiated stimulus presentation was counterbalanced across subjects. After each trial, the experimenter recorded the subject's RT and response and provided feedback concerning the response's correctness.

After a brief rest period, the subjects were given oral instructions describing the Phase 2 procedures; they were told to respond as quickly as possible without making errors. Before each test trial, the experimenter told the subject when he or she could initiate test stimulus presentation, which was initiated as in Phase 1. The stimulus presentation parameters in Phase 2 were the same as those in Phase 1, but no response feedback was given.

Subjects. Seventy-two male and female introductory psychology students received partial fulfillment of a course requirement for serving as subjects. Twenty-four subjects were assigned to each of the EPIS RGN, LDT-PNW, and LDT-UNW testing groups, according to their order of appearance in the laboratory, such that $n + 1$ subjects had not been tested in any group until n subjects had been tested in the other two groups.

Missing data and statistical analyses. Separate analyses of variance (ANOVAs) were performed on the RT and error data. For the Phase 2 data there were 10 individual subject/condition cells for which RTs had to be estimated, because a 100% error rate occurred in that cell. All of these cases occurred in the EPIS RGN test. In the EPIS RGN test, of the 192 and 144 total subject/condition cells that occurred for studied words and nonwords, respectively, RTs in three were estimated for studied words (two for one subject and one for another subject, in three different conditions) and RTs in seven were estimated for studied nonwords (two for one subject and one each for five other subjects, with four of the six conditions possessing from one to three missing cells). The data for these missing cells were estimated on an individual basis by taking into account how many standard deviations the subject's mean RT and the condition's mean RT deviated from the grand mean for the particular item type (i.e., studied word or studied nonword) in question. The inclusion of these estimated cells for the EPIS RGN test should not have biased our results greatly, for two reasons. First, fewer than 2% of the cells were estimated for studied words, which provided the data of central interest. Second, because our ANOVAs used data averaged across the correct Phase 1 answers for studied nonwords, none of the cells entering into the ANOVAs for studied nonwords was based entirely on estimated RTs.

In the ANOVAs, all factors other than subjects were fixed effects and all variables other than type of test were within-subject variables. Because the RT distributions were positively skewed, ANOVAs were performed on each subject's median RT for each condition, rather than on the mean RT. We did not treat items as a random effect (Clark, 1973), for two reasons. First, the one variable under which items were nested (i.e., frequency) had already been established as having effects that generalize over a broad range of independently sampled item sets. Second, since different items served in different levels-of-processing conditions for different subjects, the error variance due to items is already included in the subject \times condition interactions that served as the error terms for assessing the statistical significance of the various levels-of-processing effects. When variables had statistically significant interactive effects, we examined the nature of this interaction by using an *LSD* test, the error term for which was based on the *MSe* for that interaction. All effects called statistically significant were associated with two-tailed $ps < .05$.

Results and Discussion

Phase 1. Because the data of central interest come from the word stimuli in Phase 2, we will merely provide global characterizations of the word data collected in Phase 1.

For these data, the major results were that (1) semantic questions were significantly more difficult to answer than rhyme questions (a 190-msec effect for RTs and a 5.7% effect for errors); (2) low-frequency words yielded poorer performance than high-frequency words (an 83-msec effect for RTs and a 2.8% effect for errors); (3) the inferior performance on low-frequency words relative to high-frequency words was most pronounced for semantic questions for which “yes” was the correct answer. In short, the RT and error data for the words indicated that semantic questions were more difficult than rhyme questions, particularly so when subjects had to decide that an unfamiliar low-frequency word was a semantically appropriate completion to a sentence. This indicates that our levels-of-processing and word-frequency manipulations were powerful enough to affect performance during Phase 1. (For completeness, the Appendix includes the RT and error data for the various conditions in Phase 1 for both

Table 1
Means of the Median Reaction Times (RTs) and Mean Percent Errors (MPE) for the Three Groups in Phase 2 for the High-Frequency (HF) and Low-Frequency (LF) Words Presented in the Phase 1 Semantic and Rhyme Encoding Conditions

	Phase 1 Response			
	“Yes”		“No”	
	RTs	MPE	RTs	MPE
EPIS RGN				
Semantic encoding				
HF	1101	23.7	1180	37.9
LF	996	10.1	1072	37.9
WFE	-105	-13.6	-108	0.0
Rhyme encoding				
HF	1219	33.9	1119	57.1
LF	1146	31.2	1108	56.5
WFE	-73	-2.7	-11	-0.6
LDT-PNW				
Semantic encoding				
HF	600	2.3	618	2.4
LF	714	4.1	760	10.7
WFE	+114	+1.8	+142	+8.3
Rhyme encoding				
HF	648	1.7	626	2.9
LF	719	4.1	753	7.7
WFE	+71	+2.4	+127	+4.8
LDT-UNW				
Semantic encoding				
HF	571	0.0	592	2.4
LF	626	2.3	637	2.9
WFE	+55	+2.3	+45	+0.5
Rhyme encoding				
HF	590	0.0	602	0.0
LF	662	2.3	659	2.9
WFE	+72	+2.3	+57	+2.4

Note—EPIS RGN = episodic recognition; LDT-PNW = lexical decision task with pronounceable nonwords; LDT-UNW = lexical decision task with unpronounceable nonwords. WFE = word-frequency effect, which was computed by subtracting the RTs (or MPE) for HF items from the RTs (or MPE) for LF items.

Table 2
Means of the Median Reaction Times (RTs) and Mean Percent Errors (MPE) for the Three Groups in Phase 2 for the High-Frequency (HF) and Low-Frequency (LF) Words Not Presented in Phase 1

	RTs	MPE
EPIS RGN		
HF	1077	11.5
LF	1087	12.2
WFE	+10	+0.7
LDT-PNW		
HF	641	2.5
LF	824	14.5
WFE	+183	+12.0
LDT-UNW		
HF	608	0.5
LF	691	5.5
WFE	+83	+5.0

Note—EPIS RGN = episodic recognition; LDT-PNW = lexical decision task with pronounceable nonwords; LDT-UNW = lexical decision task with unpronounceable nonwords. WFE = word-frequency effect, which was computed by subtracting the RTs (or MPE) for HF items from the RTs (or MPE) for LF items.

words and nonwords, along with the results of the statistical analyses performed on those data.)

Phase 2. To answer the three questions raised in our introduction, one must compare, for the three different Phase 2 tests, the levels-of-processing effects and word-frequency effects that occurred for the words that had or had not been studied in Phase 1. Because the nonword data are not directly relevant to the main issues being addressed, they are presented in the Appendix along with a summary of the statistical analyses performed on them. Table 1 displays the RTs and errors for the words studied in Phase 1, along with the word-frequency effects they produced. Table 2 displays the same data for the nonstudied words. Table 3 presents the levels-of-processing effects obtained for studied words as a function of their frequency and of whether the correct Phase 1 response to them was “yes” or “no.”

To determine whether the levels-of-processing effects and word-frequency effects differed for the word stimuli in the three tasks, separate preliminary overall ANOVAs were performed on the RT and error data for the studied words and for the nonstudied words. Because the conditions’ median RTs and variances were positively correlated, and because EPIS RGN RTs were substantially longer than the LDT-PNW and LDT-UNW RTs, median RTs for Phase 2 were transformed by a log₁₀ transformation (see Winer, 1971, pp. 400–401). For the studied-word data, 3 (test type: EPIS RGN, LDT-PNW, and LDT-UNW) × 2 (depth of processing: rhyme vs. semantic) × 2 (correct Phase 1 answer: “yes” vs. “no”) × 2 (word frequency: high vs. low) ANOVAs were used; for the nonstudied-word data, 3 (test type: EPIS RGN, LDT-PNW, and LDT-UNW) × 2 (word frequency: high vs. low) ANOVAs were used. As shown in Tables 1 and 2, performance was worst in the EPIS RGN test, inter-

Table 3
Levels-of-Processing Effects for the Three Groups in Phase 2
for the High-Frequency (HF) and Low-Frequency (LF)
Words Presented in Phase 1

	Phase 1 Response			
	"Yes"		"No"	
	RTs	MPE	RTs	MPE
	EPIS RGN			
HF	+118	+10.2	-61	+19.2
LF	+150	+21.1	+36	+18.6
Mean	+134	+15.6	-12	+18.9
	LDT-PNW			
HF	+48	-0.6	+8	+1.2
LF	+5	0.0	-7	-2.4
Mean	+27	-0.3	0	-0.6
	LDT-UNW			
HF	+19	0.0	+10	-2.4
LF	+36	0.0	+18	0.0
Mean	+28	0.0	+14	-1.2

Note—EPIS RGN = episodic recognition; LDT-PNW = lexical decision task with pronounceable nonwords; LDT-UNW = lexical decision task with unpronounceable nonwords. The levels-of-processing effect was computed by subtracting the RTs or mean percent errors (MPE) for items that served in the semantic encoding condition in Phase 1 from the RTs or MPE for items that served in the rhyme encoding condition in Phase 1.

mediate in the LDT-PNW test, and best in the LDT-UNW test, with the main effect of test type being significant for both RTs and per cent errors for both the studied and non-studied words [all $F(2,69) > 15.9$]. But more important than the main effect of type of test are the levels-of-processing and word-frequency effects in the three different tests and their relevance to the three questions raised in our introduction, to which we now turn our attention.

(1) *Are the levels-of-processing effects similar or dissimilar in the three types of tests?* As shown in Table 3, for studied words, the levels-of-processing effects were similar for the three tests in that responses were faster [$F(1,69) = 7.67$, $MSe = .002$] and more accurate [$F(1,69) = 39.78$, $MSe = 115.81$] for words studied with semantic rather than rhyme questions. However, this levels-of-processing effect was qualified by a significant depth of processing \times correct Phase 1 answer interaction [$F(1,69) = 4.47$, $MSe = .003$] for RTs. As can be seen in Table 3, in each of the three tests for both high- and low-frequency words, the levels-of-processing effect for RTs was always larger for words given a "yes" rather than a "no" response in Phase 1. (As will be discussed in detail later, the one small levels-of-processing effect for "yes" words, i.e., the 5-msec levels-of-processing effect for low-frequency words in the LDT-PNW test, assumes importance vis à vis Forster and Davis's, 1984, explanation of the word-frequency attenuation effect.) *LSD* tests revealed that, averaged across the three tests, the +63-msec levels-of-processing effect for "yes" words was significant, whereas the +1-msec levels-of-processing effect for "no" words was not, which is a typical finding in the levels-of-processing literature (e.g., Craik and Tulving, 1975).

Although the levels-of-processing effect for RTs appeared to be larger in the EPIS RGN test than in the LDT-PNW and LDT-UNW tests, neither the test type \times depth of processing interaction nor the test type \times depth of processing \times correct Phase 1 answer interaction approached significance for RTs [both $F_s < 1.14$]. Replicating a rather typical finding in the levels-of-processing literature (e.g., Jacoby & Dallas, 1981), the present results also showed (see Table 1) that subjects made more errors for words for which the correct answer in Phase 1 had been "no," than they did for words for which the correct answer in Phase 1 had been "yes" [$F(1,69) = 83.12$, $MSe = 126.71$]. Finally, the test type \times depth of processing interaction [$F(2,69) = 47.53$, $MSe = 115.81$] and the test type \times correct Phase 1 answer interaction [$F(1,69) = 50.96$, $MSe = 126.71$] were both significant for error rates. However, this does not greatly qualify the conclusion, based on the RT data, that the levels-of-processing effects were generally similar for the three tests. Rather, since the *LSD* tests showed that the effects of depth of processing and correct Phase 1 answer were significant only in the EPIS RGN test, one can argue that the high error rates for studied words in the EPIS RGN test provided a more sensitive measure of these two variables' effects than did the very low error rates in the two LDTs.

Even if the levels-of-processing effects were larger in the EPIS RGN test than in the LDT-PNW and LDT-UNW tests, the important point is that the present data demonstrate a levels-of-processing effect in LDTs. (Additional statistical support for this conclusion is provided in the next section.) This levels-of-processing effect in an implicit recognition memory test replicates the (sometimes nonsignificant) levels-of-processing effect that Graf et al. (1982), Graf et al. (1984), Graf and Mandler (1984), Roediger et al. (1987), and Squire et al. (1987) observed in implicit word-stem completion tests (see also Note 1). This is noteworthy, because, as was noted in discussing Question 1 in the introduction, Jacoby and Dallas (1981) failed to find even a hint of a levels-of-processing effect in an implicit memory test that involved the recognition (perceptual identification) of masked words, and Kirsner et al. (1983), using a between-subjects design, obtained ambiguous levels-of-processing effects in implicit memory tests for masked stimuli in a perceptual identification task and an LDT, and for unmasked stimuli in an LDT. Thus, the present results, along with Hayman's (1983) results, demonstrate a levels-of-processing effect in an implicit memory test—that is, in an LDT—even when all of the letters of the test word are clearly presented and are not "missing" because of either visual masking or their total absence in a word-stem cue.

(2) *Do levels-of-processing effects in an LDT depend on the type of processing induced by pronounceable and unpronounceable nonword distractors?* To answer this question directly, we performed a separate 2 (test type: LDT-PNW vs. LDT-UNW) \times 2 (depth of processing: rhyme vs. semantic) \times 2 (correct Phase 1 answer: "yes" vs. "no") \times 2 (word frequency: high vs. low) ANOVA

on the RT data for the studied words in the two LDTs. (Since RTs in the two LDTs were not greatly different, a \log_{10} transformation was not applied.) The overall levels-of-processing effect was small (i.e., 18 msec), but statistically significant [$F(1,46) = 6.15$, $MSe = 5015.74$]. This shows that the significant levels-of-processing effect in the overall analysis including the EPIS RGN data was not entirely due to the numerically large levels-of-processing effect in the EPIS RGN task.

More germane to the present question is the observation that, averaged across the various kinds of studied words, the levels-of-processing effect was not smaller in the LDT-UNW test (a 22-msec effect) than in the LDT-PNW test (a 14-msec effect). That it was slightly larger calls into question data-driven/conceptually driven processing explanations for why implicit memory tests yield smaller levels-of-processing effects than do explicit memory tests. It does so by failing to support their assumption that a levels-of-processing effect will be reduced whenever a memory test, be it explicit or implicit, places greater emphasis on a test item's familiarity or perceptual fluency than on its semantic properties, as should have been so in the LDT-UNW test relative to the LDT-PNW test (e.g., see James, 1975; Shulman & Davison, 1977; Shulman, Hornak, & Sanders, 1978).

Proponents of the distinction between data-driven and conceptually driven processing could attempt to defend their approach by claiming that our specific procedures produced an artifactually inflated levels-of-processing effect in the LDT-UNW test relative to the LDT-PNW test. This counterargument hinges on the fact that the nonwords in the LDT-UNW test were not presented in their study-phase format. Because half the words in the LDT-UNW test had appeared in the study phase, whereas none of the nonwords had, the subjects could have used episodic information about a test item's study status to aid in their lexical decisions. That is, they could bias a *word* response if the test item activated episodic information that it had been studied, and they could bias a *nonword* response if it activated episodic information that it had not been studied. Since depth-of-processing has a large effect on the retrieval of episodic information, the utilization of such information in the LDT-UNW test could have artifactually inflated the levels-of-processing effect in that test. Presumably, this artifactual inflation of the levels-of-processing effect would not have occurred in the LDT-PNW test, in which the nonwords were presented in their study-phase format. That is, because half of the nonwords as well as half of the words in the LDT-PNW test had appeared in the study phase, episodic information concerning a test item's study status would not be as valid a predictor of a test item's lexical status in the LDT-PNW test. Thus, the equivalent levels-of-processing effects in the two LDTs could have been due to an artifactual enhancement of the levels-of-processing effect in the LDT-UNW task, in which episodic information could be used to enhance performance, relative to the LDT-PNW task, in which it could not be as validly used.

However, there are at least two problems with this episodic response-bias explanation of the equivalent levels-of-processing effects that occurred in the two LDTs. First, this explanation must incorrectly assume that studied words and nonwords had equally strong episodic memory traces at the time the LDT-PNW test was administered. (That this assumption is incorrect is demonstrated by the very high error rates that occurred for nonwords in the EPIS RGN test—see Table 2 in the Appendix.) Since words were remembered better episodically than were nonwords, the subjects could have used episodic information that an item had been studied in order to bias a *word* response in both LDTs. To the degree that they did, a differential utilization of episodic information in these two tests cannot account for their equivalent levels-of-processing effects.

A second more severe problem with the episodic response-bias account of the equivalent levels-of-processing effects that occurred in the two LDTs is that it incorrectly predicts that the effect of repetition priming from prior presentation in the study phase should be larger in the LDT-UNW test than in the LDT-PNW test. Because subjects in the LDT-UNW test presumably could use episodic information better in order to bias their "word" responses for words presented in the study phase than could subjects in the LDT-PNW test, the subjects in the LDT-UNW test would be more biased than subjects in the LDT-PNW test to respond *word* to studied words relative to nonstudied words, thereby enhancing their repetition priming effects. Because repetition priming effects were larger, not smaller, in the LDT-PNW test than in the LDT-UNW test, we tend to discount the episodic response bias explanation and believe that the present data indeed cause problems for data-driven/conceptually driven processing accounts of why dissociative levels-of-processing effects occur for implicit as opposed to explicit memory tests.

The equivalence of the levels-of-processing effects in the LDT-UNW and LDT-PNW tests is important for another reason. This equivalence constrains the interpretation of Hayman's (1983) finding that a levels-of-processing effect is obtained in an LDT-PNW test when the instructions emphasize accuracy but not when they emphasize speed. The present results suggest that Hayman's (1983) finding may not have been due to processing speed per se, since the presence of unpronounceable nonword distractors sped up *word* lexical decisions by 63 msec—though this between-subjects effect for the studied words was not significant [$F(1,46) = 1.75$, $MSe = 30,852.51$]—but did not reduce the levels-of-processing effect observed for them. One tentative conclusion from this is that levels-of-processing effects are reduced in an LDT if one manipulates a variable that speeds up RTs and decreases accuracy (Hayman, 1983), but not if one manipulates a variable that speeds up RTs without decreasing accuracy (the present results).

(3) *Is explicit episodic memory responsible for repetition's attenuating the word-frequency effect in the LDT?*

To answer this question, we must first determine if we obtained typical word-frequency effects in our three tasks—that is, better performance on low- than on high-frequency words in the EPIS RGN test and better performance on high- than low-frequency words in the two LDTs, with the latter effect being attenuated for studied words relative to nonstudied words. As shown in Tables 1 and 2, the word-frequency effects for both studied and nonstudied words were quite different in the three tests. This conclusion was supported by significant test type \times word frequency interactions for both RTs and errors for both studied and nonstudied words [all $F_s(2,69) > 7.5$]. *LSD* tests were used to examine the word-frequency effects for the three tests. Averaged across cells representing differences in processing depth and correct Phase 1 answer, the results were as follows: for studied words (see Table 1), word-frequency effects of -74 msec (-4.2%), $+113$ msec ($+4.3\%$), and $+57$ msec ($+2.0\%$), in the EPIS RGN, LDT-PNW, and LDT-UNW tests, respectively, were all significant, apart from the $+1.9\%$ effect for errors in the LDT-UNW test; for nonstudied words (see Table 2), the $+10$ -msec and $+0.7\%$ word-frequency effects in the EPIS RGN test were not significant, whereas the $+183$ -msec and $+12.0\%$ word-frequency effects in the LDT-PNW test and the $+83$ msec and $+5.0\%$ word-frequency effects in the LDT-UNW test were all significant. That the word-frequency effects in the EPIS RGN test did not show the same patterns for the studied words (which required a “yes” response) and the nonstudied words (which required a “no” response) indicates that the word-frequency effects for studied words were due to a true memory difference and not to some sort of response bias effect (cf. Glanzer & Adams, 1985; Glanzer & Bowles, 1976).

These word-frequency effects clearly replicate those from the many previous studies we have cited in our introduction. That is, we too found better performance for low- than for high-frequency words in our EPIS RGN test, whereas we found the opposite in our two LDTs. The importance of the present results was that these dissociative word-frequency effects for episodic recognition and LDTs were obtained under conditions in which these two types of memory tasks were equated in terms of (1) the test items’ having received a prior study presentation, (2) the type of processing that the subjects performed during this prior study presentation, and (3) the particular high- and low-frequency words and nonwords that served as test items. These controls are important (see Neely, 1989), because the word-frequency effects that occur within a lexical decision or an episodic recognition test are influenced by these three variables. Specifically, (1) the word-frequency effect for lexical decisions is reduced when the high- and low-frequency words have received prior study (Forster & Davis, 1984; Scarborough et al., 1977); (2) the magnitude of the word-frequency effect in episodic recognition depends on how the test words were encoded during study (e.g., see Balota & Neely, 1980; Eysenck, 1979; Mandler, Goodman, & Wilkes-Gibbs, 1982); and (3) the word-frequency effect in the LDT is

influenced by the type of nonword distractor (pronounceable vs. unpronounceable) that is used (James, 1975) and by the proportion of high- and low-frequency words appearing on the test list (Glanzer & Ehrenreich, 1979; Gordon, 1983; Neely, 1980), whereas the word-frequency effect in episodic recognition is reversed for very low-frequency words (Mandler et al., 1982; Schulman, 1976). Because we controlled these variables, as well as all other variables apart from the instructions, the dissociative word-frequency effects we obtained in our EPIS RGN test and our two LDTs were due to differences in the types of information retrieved in the two types of test, and not to differences in other confounded variables.

To determine whether repetition priming from prior study attenuated the word-frequency effects obtained for lexical decisions, we submitted the (untransformed) median RT and error data for lexical decisions to words to separate 2 (test type: LDT-PNW vs. LDT-UNW) \times 2 (word frequency: high vs. low) \times 5 (item type: semantic “yes,” semantic “no,” rhyme “yes,” rhyme “no,” vs. nonstudied) ANOVAs. For both RTs and errors, the word-frequency effects were highly significant [both $F_s(1,46) > 76.30$], and, replicating results reported by James (1975), they were larger in the LDT-PNW test (127 msec and 7.2%) than in the LDT-UNW test (62 msec and 2.6%). Both $F_s(1,46) > 17.10$ for the word-frequency \times test type interaction. More importantly, the item type \times word frequency interaction was significant both for RTs [$F(4,179) = 3.04$, $MSe = 4181.17$], and for errors [$F(4,179) = 3.13$, $MSe = 42.33$]. Specifically, averaged across the LDT-PNW and LDT-UNW tests, the *LSD* tests showed that the word-frequency effects obtained for RTs for the four types of studied items (which ranged from 72–94 msec; see Table 1) were all significantly smaller than the 133-msec word-frequency effect obtained for the nonstudied items (see Table 2). (Although the word-frequency effects for errors for studied items were all smaller than those for the nonstudied items, they were significantly smaller only for the items for which “yes” had been the correct Phase 1 answer.) Finally, the test type \times word frequency \times item type interaction was of borderline significance for RTs [$F(4,184) = 2.41$, $MSe = 4181.17$, $p < .052$], and it reached conventional levels of significance for errors [$F(4,184) = 2.73$, $MSe = 42.33$]. The nature of this interaction was that the word-frequency attenuation effect from prior study was larger in the LDT-PNW test than in the LDT-UNW test. That is, in the LDT-PNW test, the word-frequency effects were 114 msec and 7.3% averaged across the four different types of studied items (see Table 1), and they were 183 msec and 12.0% for the nonstudied items (see Table 2), which yielded word-frequency attenuation effects of 69 msec and 4.7%; in the LDT-UNW test, the word-frequency effects were 57 msec and 2.0% for the studied items (see Table 1), and 83 msec and 5.0% for nonstudied items (see Table 2), which yielded word-frequency attenuation effects of only 26 msec and 3.0%. In short, the present data clearly demonstrate a substantial word-frequency attenuation ef-

fect in the LDT-PNW test, but only a modest word-frequency attenuation effect in the LDT-UNW test.

Clearly, the word-frequency attenuation effects produced by repetition priming in the LDT-PNW test replicate results reported by Forster and Davis (1984) and Scarborough et al. (1977). Thus, we can test the generality of Forster and Davis's (1984) claim that word-frequency attenuation effects are due to lexical decisions for low-frequency words' receiving an episodic mnemonic benefit from their prior study greater than that received by lexical decisions for high-frequency words. To test this directly, we need to compare how our depth-of-processing manipulation affected performance in response to high- and low-frequency studied words in the EPIS RGN and the LDT-PNW tests. A 2 (test type: EPIS RGN vs. LDT-PNW) \times 2 (depth of processing: semantic vs. rhyme) \times 2 (correct Phase 1 answer: "yes" vs. "no") \times 2 (word frequency: high vs. low) ANOVA performed on the \log_{10} median RTs revealed a significant test type \times depth of processing \times word frequency interaction [$F(1,46) = 4.21$, $MSe = .0024$]. The *LSD* tests showed that, averaged across the correct Phase 1 answer, in the EPIS RGN test (see Table 3) the +94-msec levels-of-processing effect for low-frequency words was significant, whereas the +28-msec levels-of-processing effect for high-frequency words was not. This suggests that "deeper" semantic processing benefitted episodic memory more for low- than for high-frequency words. Thus, if Forster and Davis' (1984) claim that episodic memory mediates the word-frequency attenuation effect for lexical decisions is true, "deeper" semantic processing should have also benefitted RTs for studied low-frequency words more than for studied high-frequency words in the LDT-PNW test. But exactly the opposite was so. That is, in the LDT-PNW test (see Table 3), the -1-msec levels-of-processing effect for low-frequency studied words was smaller than the statistically significant +28-msec levels-of-processing effect for high-frequency words.

Although the foregoing argument seems weakened by the fact that the +28-msec levels-of-processing effect was significant for the LDT-PNW test but not for the EPIS RGN test, this fact does not vitiate our argument. This difference in significance levels is due to the \log_{10} transformation's compressing the differences based on the much longer (and more highly variable) RTs in the EPIS RGN test. More importantly, the conclusion that levels-of-processing had different effects on RTs with respect to low- as opposed to high-frequency words in the EPIS RGN and LDT-PNW tests is also supported by the fact that the +66 msec (i.e., +94 minus +28 msec) difference between the levels-of-processing effects obtained for low- and high-frequency words in the EPIS RGN test is clearly significantly different from the -29-msec (i.e., -1 minus +28 msec) difference observed in the LDT-PNW test, even if one chooses to believe the latter difference was not statistically significant.

In short, these data contradict Forster and Davis's (1984) argument that the word-frequency attenuation ef-

fect for lexical decisions produced by prior presentations of the high- and low-frequency words is mediated by episodic memory for these prior presentations. However, this contradiction is not all that surprising, given Moscovitch's (1982) and Scarborough et al.'s (1977) data, cited in our introduction. Their data suggest that even when the repeated item's first presentation is unmasked and attended, episodic memory may not in any way be involved in producing repetition priming effects in the LDT.

GENERAL DISCUSSION

The major empirical thrust of the present research is embodied in three main findings. The first main finding was that a statistically significant, albeit numerically small, levels-of-processing effect occurred in an implicit memory test. Specifically, we found that implicit memory for the prior presentation of a word (as measured by the amount of repetition priming that prior presentation produced in an LDT) was greater when subjects had processed that prior presentation by answering a "deep" semantic question about it rather than a "shallow" rhyme question. It is also interesting that, as in our explicit recognition memory test (see also Craik & Tulving, 1975), this levels-of-processing effect in implicit memory was larger when the correct answer to the semantic and rhyme questions had been "yes." Although this suggests that the levels-of-processing effects obtained in explicit and implicit memory tests are similar, they may also differ in some ways. For example, as was observed here, levels-of-processing effects are always numerically (though not always statistically) larger in explicit than in implicit memory tests (Graf & Mandler, 1984; Graf et al., 1982; Graf et al., 1984; Jacoby & Dallas, 1981; Kirsner et al., 1983; Roediger et al., 1987; Squire et al., 1987). However, even if levels-of-processing effects are larger in explicit than in implicit memory tests, the present results and Hayman's (1983) results clearly demonstrate that statistically significant levels-of-processing effects can be obtained when one measures implicit memory by examining repetition priming effects in an LDT. Such a demonstration is contrary to the conclusion strongly implied by many (e.g., Graf & Mandler, 1984; Graf et al., 1982; Graf et al., 1984; Jacoby & Dallas, 1981; Kirsner et al., 1983) that levels-of-processing effects do not occur in implicit memory tests.

The second major finding of the present research is that the levels-of-processing effect was just as large in our LDT-UNW task as in our LDT-PNW task. This finding is interesting because, as was discussed at the end of the Results section with regard to Question 2, it shows that speed of processing per se does not modulate levels-of-processing effects. More importantly, this finding is problematic for the widely accepted data-driven/conceptually driven processing account for why levels-of-processing effects are smaller in implicit than in explicit memory tasks (e.g., Blaxton, 1985; Graf & Mandler, 1984; Jacoby, 1983; Jacoby & Dallas, 1981; Roediger

& Blaxton, 1987; Roediger et al., 1987). According to this account, implicit memory tests typically invoke data-driven processing operations that tap an item's perceptual fluency or familiarity, whereas explicit memory tests typically invoke conceptually driven processing operations that tap information about an item's elaborative associations with other studied items, with the study context, and with other related items in semantic memory. Since prior research (see Balota & Chumbley, 1984; James, 1975; Shulman & Davison, 1977; Shulman, Hornak, & Sanders, 1978) has suggested that data-driven processing is more likely to be utilized in the LDT-UNW test than in the LDT-PNW test, an analysis based on data-driven and conceptually driven processing predicts that levels-of-processing effects should be smaller in an LDT-UNW test than in an LDT-PNW test. That the levels-of-processing effects in these two tests were equivalent therefore creates problems for this analysis.

The third main finding of the present research is that word-frequency effects were dissociated in explicit and implicit memory tests. These dissociative word-frequency effects took two forms. First, low-frequency words yielded better performance than high-frequency words in our explicit EPIS RGN test, whereas the opposite was so in our LDT-PNW test. Although these two opposite effects have been obtained across different experiments with different materials, encoding conditions, etc., as far as we know we are the first to demonstrate them with implicit and explicit memory tasks differing only in the instructions that the subjects received as to how they were to perform them. Second, in the EPIS RGN test, memory for low-frequency words was helped more by prior semantic processing than was memory for high-frequency words, whereas in the LDT-PNW test the opposite was so. As we argued in detail in the Results section with regard to Question 3, this latter dissociation causes problems for Forster and Davis's (1984) claim that episodic familiarity mediates the attenuation of the word-frequency effect that occurs in implicit memory tests when the high- and low-frequency words have been previously presented. However, we must acknowledge two qualifications to our argument.

The first qualification is that we are not arguing that episodic memory *never* plays a role in the word-frequency attenuation effect in implicit memory tasks. We are merely arguing that (1) prior experiments on this issue have provided no *direct* evidence in favor of Forster and Davis's (1984) claim that episodic memory is indeed involved in the word-frequency attenuation effect, and (2) *under the conditions used in the present experiment*, the evidence actually stands in direct conflict with that claim. The second qualification concerns the generality of the word-frequency \times levels-of-processing interaction that we obtained in our EPIS RGN test. Mandler et al. (1982, Experiment 2) manipulated depth of processing in a fashion analogous to that used here, but between, rather than within, subjects. With percent hits – percent false alarms as a performance measure, Mandler et al. (1982) found that high-frequency words yielded a larger levels-

of-processing effect than did low-frequency words, which is opposite to the result that we obtained. However, our levels-of-processing effect occurred only in hit rates (it could not occur in false-alarm rates in our within-subject manipulation of depth of processing), whereas in Mandler et al.'s data it occurred only in false alarm rates. Although it is at present unclear which of the many procedural differences between Mandler et al.'s and our experiments produced these discrepancies in results, the discrepancies do not undermine our claim that the present data argue against the involvement of episodic memory in the word-frequency attenuation effect. The reason they do not is that we examined the effect that levels-of-processing had on episodic memory for the high- and low-frequency words under exactly the same conditions as those used to examine the word-frequency attenuation effect in the LDT-PNW test.

CONCLUSION

The present results clearly demonstrate that (1) a levels-of-processing effect can be obtained in at least one implicit memory test (i.e., the LDT); (2) the magnitude of this levels-of-processing effect is not affected by whether a "deep" or a "shallow" level of analysis is necessary to make the lexical decisions; (3) the effect that levels-of-processing has on memory for high- and low-frequency words is different for explicit and implicit memory tests. These results pose problems for the view (e.g., Blaxton, 1985; Graf & Mandler, 1984; Jacoby, 1983; Jacoby & Dallas, 1981; Roediger & Blaxton, 1987; Roediger et al., 1987) that levels-of-processing effects are smaller for implicit than for explicit memory tests, because the former tests typically involve data-driven processing whereas the latter tests typically involve conceptually driven processing. They also pose problems for the view (e.g., Forster & Davis, 1984) that explicit episodic memory plays an important role in the word-frequency attenuation effect in the LDT.

Exactly how these views can be modified to accommodate the present results remains unclear. What is clear, as Roediger (1984) has warned, is that it is dangerous to generalize results from one implicit (or explicit) memory task to another task of the same general type. On the basis of the present results, the next order of business is to isolate the processing mechanisms unique to the LDT that result in its yielding a levels-of-processing effect while other implicit memory measures do not. We hope that the present results can serve as a springboard for such an enterprise.

REFERENCES

- BALOTA, D. A. (1983). Automatic semantic activation and episodic memory encoding. *Journal of Verbal Learning & Verbal Behavior*, 22, 88-104.
- BALOTA, D. A., & CHUMBLEY, J. I. (1984). Are lexical decisions a good measure of lexical access: The role of word frequency in the neglected decision stage? *Journal of Experimental Psychology: Human Perception & Performance*, 10, 340-357.
- BALOTA, D. A., & NEELY, J. H. (1980). Test-expectancy and word-

- frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 576-587.
- BESNER, D., & SWAN, M. (1982). Models of lexical access in visual word recognition. *Quarterly Journal of Experimental Psychology*, **34A**, 313-325.
- BLAXTON, T. A. (1985). *Examining theoretical accounts of dissociations among memory measures*. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.
- CLARK, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior*, **12**, 335-359.
- CRAIK, F. I. M., & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, **11**, 671-684.
- 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.
- DURGUNOĞLU, A., & NEELY, J. H. (1987). On obtaining episodic priming in a lexical decision task following paired-associate learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 206-222.
- EYSENCK, M. W. (1979). Depth, elaboration, and distinctiveness. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 89-118). Hillsdale, NJ: Erlbaum.
- FORSTER, K. I., & DAVIS, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 680-698.
- GLANZER, M., & ADAMS, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, **13**, 8-20.
- GLANZER, M., & BOWLES, N. (1976). Analysis of the word frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 21-31.
- GLANZER, M., & EHRENREICH, S. L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning & Verbal Behavior*, **18**, 381-398.
- GORDON, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning & Verbal Behavior*, **22**, 24-44.
- GRAF, P., & MANDLER, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning & Verbal Behavior*, **23**, 553-568.
- GRAF, P., MANDLER, G., & HADEN, P. (1982). Simulating amnesic symptoms in normal subjects. *Science*, **218**, 1243-1244.
- GRAF, P., & SCHACTER, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 501-518.
- GRAF, P., SQUIRE, L. R., & MANDLER, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 164-178.
- GREGG, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). New York: Wiley.
- HAYMAN, C. A. G. (1983). *A task analysis of lexical decisions*. Unpublished doctoral dissertation, McMaster University, Ontario, Canada.
- JACOBY, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, **22**, 485-508.
- JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.
- JAMES, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 130-136.
- JOHNSTON, W. A., DARK, V. J., & JACOBY, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 3-11.
- KIRSNER, K., MILECH, D., & STANDEN, P. (1983). Common and modality-specific processes in the mental lexicon. *Memory & Cognition*, **11**, 621-630.
- KIRSNER, K., & SMITH, M. C. (1974). Modality effects in word identification. *Memory & Cognition*, **2**, 637-640.
- KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.
- LOGAN, G. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.
- MANDLER, G., GOODMAN, G. O., & WILKES-GIBBS, D. L. (1982). The word-frequency paradox in recognition. *Memory & Cognition*, **10**, 33-42.
- MCKOON, G., & RATCLIFF, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning & Verbal Behavior*, **18**, 463-480.
- MITTERER, J. O., & BEGG, I. (1979). Can meaning be extracted from meaningless stimuli? *Canadian Journal of Psychology*, **33**, 193-197.
- MONSELL, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp. 147-195). London: Erlbaum.
- MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.
- MOSCOVITCH, M. (1982). A neuropsychological approach to perception and memory in normal and pathological aging. In F. I. M. Craik & S. Trehub (Eds.), *Memory and cognitive processes in aging* (pp. 55-79). New York: Plenum Press.
- NEELY, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**, 226-254.
- NEELY, J. H. (1980, May). *The role of expectancy in the word-frequency effect in a lexical decision task*. Paper presented at the meeting of the Midwestern Psychological Association, Chicago.
- NEELY, J. H. (1989). Experimental dissociations and the episodic/semantic memory distinction. In H. L. Roediger III and F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 229-270). Hillsdale, NJ: Erlbaum.
- NEELY, J. H., & DURGUNOĞLU, A. (1985). Dissociative episodic and semantic priming effects in episodic recognition and lexical decision tasks. *Journal of Memory & Language*, **24**, 466-489.
- NEELY, J. H., & PAYNE, D. G. (1983). A direct comparison of recognition failure rates for recallable names in episodic and semantic memory tests. *Memory & Cognition*, **11**, 161-171.
- NORRIS, D. (1984). The effects of frequency, repetition and stimulus quality in visual word recognition. *Quarterly Journal of Experimental Psychology*, **36A**, 507-518.
- RICHARDSON-KLAVEHN, A., & BJORK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475-543.
- ROEDIGER, H. L. (1984). Does current evidence from dissociation experiments favor the episodic/semantic memory distinction? *Behavioral & Brain Sciences*, **7**, 252-254.
- ROEDIGER, H. L., & 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* (pp. 349-379). Hillsdale, NJ: Erlbaum.
- ROEDIGER, H. L., WELDON, M. S., & STADLER, M. A. (1987, November). *Direct comparison of two implicit measures of retention*. Paper presented at the meeting of the Psychonomic Society, Seattle, WA.
- SCARBOROUGH, D. L., CORTESE, C., & SCARBOROUGH, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 1-17.
- SCHACTER, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 501-518.
- SCHULMAN, A. I. (1976). Memory for rare words previously rated for familiarity. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 301-307.
- SHEPARD, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior*, **6**, 156-163.
- SHULMAN, H. G., & DAVISON, T. C. B. (1977). Control properties of

- semantic coding in a lexical decision task. *Journal of Verbal Learning & Verbal Behavior*, 16, 92-98.
- SHULMAN, H. G., HORNACK, R., & SANDERS, E. (1978). The effects of graphemic, phonetic, and semantic relationships on access to lexical structures. *Memory & Cognition*, 6, 115-123.
- SQUIRE, L. R., SHIMAMURA, A. P., & GRAF, P. (1987). Strength and duration of priming effects in normal subjects and amnesic patients. *Neuropsychologia*, 25, 195-210.
- STANNERS, R. F., JASTRZEMBSKI, J. E., & WESTBROOK, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning & Verbal Behavior*, 14, 259-264.
- TULVING, E. (1985). How many memory systems are there? *American Psychologist*, 40, 385-398.
- WINER, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.
- WOLFORD, G., & MORRISON, F. (1980). Processing of unattended visual information. *Memory & Cognition*, 8, 521-527.

NOTE

1. In the Graf et al. (1984) experiment, the levels-of-processing effect that occurred in the implicit memory test could have been a carry-over from the levels-of-processing effect that occurred in the retrieval of the same test items in a preceding explicit memory test. Also, although Squire et al. (1987) obtained statistically significant levels-of-processing effects in both word-stem and word-fragment completion tests, in which letters other than the initial letters of the word served as cues, Roediger et al. (1987) found a nonsignificant levels-of-processing effect in their word-stem completion test and a nonsignificant reversal of the normal levels-of-processing effect in their word-fragment completion test. It is unclear why this discrepancy in results occurred, though it should be noted that Roediger et al. used exactly the same items as targets in their word-stem and word-completion tests, whereas Squire et al. did not.

APPENDIX

For the sake of completeness, this Appendix summarizes the statistical analyses performed on the data not directly relevant to the three main questions addressed in the text of the preceding article.

Phase 1 Data

Table A1 summarizes the means of the median RTs and the error data for the various conditions in Phase 1. Because all subjects were treated identically in Phase 1, these data are averaged across the subjects who would be receiving the three different Phase 2 tests. Also, because "no" was always the correct response for nonwords in the semantic encoding condition, the

Table A1
Means of the Median Reaction Times (RTs) and Mean Percent Errors (MPE) in Phase 1 for the High-Frequency (HF) and Low-Frequency (LF) Items in the Semantic and Rhyme Encoding Conditions

	Item Type					
	"Yes"-Word		"No"-Word		Nonword*	
	RTs	MPE	RTs	MPE	RTs	MPE
Semantic Encoding						
HF	892	5.1	1012	3.5	952	1.2
LF	1053	14.9	1065	4.5	1046	2.6
Rhyme Encoding						
HF	768	2.0	806	1.6	900	3.9
LF	842	0.6	850	1.3	919	3.4

*The nonword data are averaged across the "yes" and "no" answers for the rhyme encoding condition.

nonword data for the rhyme encoding condition are presented as averaged across the "yes" and "no" responses. Separate ANOVAs were performed on the median RTs for correct responses and on the percentage errors for the word stimuli and for the nonword stimuli. The word data were analyzed in 2 (depth of processing: rhyme vs. semantic) \times 2 (correct answer: "yes" vs. "no") \times 2 (word frequency: high vs. low) ANOVAs, and the nonword data were analyzed in 2 (depth of processing: semantic vs. rhyme) \times 2 (frequency: high vs. low) ANOVAs.

For words, the word-frequency effect was significant both for RTs [$F(1,69) = 32.64$, $MSe = 30352.19$] and for errors [$F(1,69) = 16.69$, $MSe = 66.24$], with responses being 83 msec faster and 2.8% more accurate for high- than for low-frequency words. The levels-of-processing effect was also significant for the words, both for RTs [$F(1,69) = 99.82$, $MSe = 51536.25$] and for errors [$F(1,69) = 62.55$, $MSe = 73.74$], with RTs being 190 msec faster and 5.7% more accurate for the rhyme questions than for the semantic questions. Furthermore, the word frequency \times correct answer interaction was significant for both RTs [$F(1,69) = 12.52$, $MSe = 13862.67$] and errors [$F(1,69) = 12.22$, $MSe = 43.59$], with the word-frequency effect being larger for "yes" answers (117 msec and 4.2%) than for "no" answers (48 msec and .3%). Although there were no other significant effects for RTs, there were several other significant effects for error rates. First, as shown in Table A1, fewer errors were made for "no" answers (2.5%) than for "yes" answers (5.9%) [$F(1,69) = 20.92$, $MSe = 78.42$]. Second, the depth of processing \times correct answer [$F(1,69) = 16.06$, $MSe = 58.25$] and depth of processing \times word frequency [$F(1,69) = 14.75$, $MSe = 66.24$] interactions were both significant. Specifically, the levels-of-processing effect was greater for "yes" answers (8.7%) than for "no" answers (2.5%), and, averaged across "yes" and "no" answers, it was greater for low than for high frequency words (8.8% vs. 2.5%). However, these two interactions were themselves qualified by a significant depth of processing \times word frequency \times correct answer interaction [$F(1,69) = 24.24$, $MSe = 36.31$], which was due to the 14.3% levels-of-processing effect for "yes" low-frequency words' being much greater than the 1.9%-3.2% levels-of-processing effects obtained for "no" low-frequency words or "yes" or "no" high-frequency words.

As shown in the right-hand column of Table A1, for nonwords there was a 90 msec levels-of-processing effect [$F(1,69) = 8.68$, $MSe = 66046.64$] and a 56-msec frequency effect [$F(1,69) = 32.64$, $MSe = 30352.19$], with RTs being faster for "high-" than for "low-frequency" nonwords, and faster for rhyme than for semantic questions. However, there was also a significant depth of processing \times frequency interaction [$F(1,69) = 13.62$, $MSe = 7520.63$], with frequency having a large (statistically significant) 94-msec effect for semantic questions, but only a small (nonsignificant) 19-msec effect for rhyme questions. The error rates, which were very low, yielded no significant effects.

Phase 2 Data

Nonword data. Tables A2 and A3 present the Phase 2 data for the studied and nonstudied nonwords, respectively. Once again, the studied nonword data in the rhyme encoding condition are averaged across "yes" and "no" responses. A 3 (test type: EPIS RGN, LDT-PNW, and LDT-UNW) \times 2 (depth of processing: rhyme vs. semantic) \times 2 (frequency: high vs. low) ANOVA was used to analyze the data for the studied nonwords, and a 3 (test type: EPIS RGN, LDT-PNW, and LDT-UNW) \times 2 (frequency: high vs. low) ANOVA was used for the non-studied nonword data. A \log_{10} transformation was applied to the median RT for each condition for each subject.

Table A2
Means of the Median Reaction Times (RTs) and Mean Percent Errors (MPE) for the Three Groups in Phase 2 for the High-Frequency (HF) and Low-Frequency (LF) Nonwords Presented in the Phase 1 Semantic and Rhyme Encoding Conditions

	HF	LF	WFE		
EPIS RGN					
Semantic encoding	1188	62.5	1332	48.9	+144 -13.6
Rhyme encoding	1230	52.3	1216	57.2	-14 +4.9
LOPE	+42	-10.2	-116	+8.3	
LDT-PNW					
Semantic encoding	809	2.9	823	9.1	+14 +6.2
Rhyme encoding	853	6.7	878	8.6	+25 +1.9
LOPE	+44	+3.8	+55	-0.5	
LDT-UNW					
Semantic encoding	660	1.2	664	1.5	+4 +0.3
Rhyme encoding	639	0.3	680	2.7	+41 +2.4
LOPE	-21	-0.9	+16	+1.2	

Note—EPIS RGN = episodic recognition; LDT-PNW = lexical decision task with pronounceable nonwords; LDT-UNW = lexical decision task with unpronounceable nonwords. WFE = word-frequency effect, which was computed by subtracting the RTs (or MPE) for HF items from the RTs (or MPE) for LF items. LOPE = levels-of-processing effect, which was computed by subtracting the RTs or mean percent errors (MPE) for the Semantic Encoding condition from the RTs or MPE for the Rhyme Encoding condition.

RTs were faster for both studied and nonstudied nonwords [both $F(1,69) > 9.9$], and more accurate for nonstudied nonwords [$F(1,69) = 11.26, MSe = 10.74$], when these nonwords were derived from high- rather than low-frequency words. This seemingly curious “frequency” effect for nonwords has also been reported by Neely (1977) and Stanners, Jastrzembski, and Westbrook (1975). Also, as in the present experiment, their frequency effects were larger for words than for nonwords. For RTs for studied nonwords there was also a significant test type × depth of processing interaction [$F(2,69) = 6.93, MSe =$

Table A3
Means of the Median Reaction Times (RTs) and Mean Percent Errors (MPE) for the Three Groups in Phase 2 for the High-Frequency (HF) and Low-Frequency (LF) Nonwords Not Presented in Phase 1

	RTs	MPE
EPIS RGN		
HF	1052	7.8
LF	1074	10.4
WFE	+22	+2.6
LDT-PNW		
HF	808	2.8
LF	853	4.8
WFE	+45	+2.0
LDT-UNW		
HF	642	1.0
LF	654	1.9
WFE	+12	+0.9

Note—EPIS RGN = episodic recognition; LDT-PNW = lexical decision task with pronounceable nonwords; LDT-UNW = lexical decision task with unpronounceable nonwords. WFE = word-frequency effect, which was computed by subtracting the RTs (or MPE) for HF items from the RTs (or MPE) for LF items.

.001]. As shown in Table A2, the levels-of-processing effect for nonwords was different for the EPIS RGN and the LDT-PNW tests, in which the test format of the nonwords was identical to their study format. That is, in the LDT-PNW test, “deeper” semantic processing facilitated RTs with respect to studied nonwords by 49 msec, whereas in the EPIS RGN test, “deeper” semantic processing inhibited RTs to studied nonwords by 37 msec. (The *LSD* tests revealed that the levels-of-processing effect for nonwords was significant only in the LDT-PNW test.) Although somewhat curious, a levels-of-processing effect for nonwords is not without precedent. Specifically, Mitterer and Begg (1979) have reported a levels-of-processing effect for both pronounceable and unpronounceable nonwords, but their effect occurred in an episodic recognition test; we failed to find that effect here.

The only other significant effects occurred for errors in responses to studied nonwords. Specifically, the test type × frequency interaction [$F(2,69) = 7.26, MSe = 59.87$] and the frequency × depth of processing interaction [$F(1,69) = 9.40, MSe = 56.75$] were both significant for errors in responses to studied nonwords. However, these two interactions were further qualified by a significant test type × frequency × depth of processing interaction [$F(2,69) = 14.78, MSe = 56.75$]. The *LSD* tests showed that the following effects were significant: In the EPIS RGN test, the subjects made 13.6% more errors on “high-” than on “low-frequency” nonwords studied with semantic questions, whereas for nonwords studied with rhyme questions, they made 4.9% fewer errors; in the LDT-PNW test, on the other hand, when the nonwords had been studied with semantic questions, the subjects made 6.2% fewer errors on high- than on “low-frequency” nonwords. The remaining frequency effects for errors in responses to studied nonwords (i.e., those for rhyme-encoded nonwords in the LDT-PNW test and for rhyme-encoded and semantically encoded nonwords in the LDT-UNW test) were not statistically significant.

Effects of prior study on word and nonword processing in the LDT-PNW test. In the final ANOVAs we report, we examined how prior study affected RTs and error rates in responses to word and nonword stimuli in the LDT-PNW test. To do so, we averaged the data across the different types of studied items and performed 2 (prior study: studied vs. nonstudied) × 2 (lexicality: word vs. nonword) × 2 (frequency: high vs. low) ANOVAs on the RT and error data. These ANOVAs indicated that for RTs all main effects and all interactions were significant, and that for errors all but the main effects of prior study and lexicality, as well as all interactions but the prior study × frequency interactions, were significant.

It is particularly interesting that prior study facilitated *word* responses but inhibited *nonword* responses, with the prior study × lexicality interactions yielding $F(1,23) = 19.87, MSe = 3088.48$, for RTs, and $F(1,23) = 21.76, MSe = 30.22$, for errors. Specifically, the *LSD* tests showed that (1) the 53-msec and 4.0% facilitation effects that prior study produced for *word* RTs and error rates were both significant, and (2) of the 19-msec and 3.4% inhibition effects that prior study produced for *nonword* responses, the inhibition for errors was significant, but the inhibition for RTs was only marginally reliable ($p < .055$, one-tailed). However, the differential effect that prior study had on *word* and *nonword* responses was modulated by the frequency variable, which was indicated by the significant prior study × lexicality × frequency interactions: $F(1,23) = 4.35, MSe = 1468.38$, for RTs; $F(1,23) = 7.54, MSe = 14.73$, for errors. The *LSD* tests indicated that the nature of this interaction for RTs was that for the facilitation effects that prior study produced

for *word* responses, the 18-msec facilitation effect for high-frequency words was not significant, whereas the 87-msec facilitation effect for low-frequency words was. This, of course, is merely the word-frequency attenuation effect produced by prior study. On the other hand, for the inhibition effects that prior study produced for *nonword* responses, the 30-msec inhibition for "high-frequency" nonwords was significant, whereas the 7-msec inhibition effect for "low-frequency" nonwords was not. (For errors, the *LSD* tests showed that of the four facilitation and inhibition effects produced by prior study, all but the facilitation effect for high-frequency words was significant.)

Though the inhibition that prior study produced for *nonword* responses runs counter to Besner and Swan's (1982), Kirsner and Smith's (1974), Logan's (1988), Monsell's (1985), Norris' (1984), and Scarborough et al.'s (1977) findings that repetition facilitates both *word* and *nonword* responses in the LDT, it replicates results reported by Durgunoğlu and Neely (1987), Hayman (1983), McKoon and Ratcliff (1979), and Neely and Durgunoğlu (1985), all of whom showed that prior study facilitated *word* responses and inhibited *nonword* responses in their LDTs.

The variable that most likely accounts for whether repetition will produce facilitation or inhibition for *nonword* responses in the LDT is the kind of encoding done on the repeated nonword's first presentation: if it was a lexical decision, facilitation occurs (Besner & Swan, 1982; Kirsner & Smith, 1974; Logan, 1988; Norris, 1984; Scarborough et al., 1977); if it was paired-associate study in anticipation of a cued recall test (Durgunoğlu & Neely, 1987; McKoon & Ratcliff, 1979; Neely & Durgunoğlu, 1985) or answering rhyme or semantic questions about the nonword (Hayman, 1983, and the present experiment), inhibition occurs. The one exception to this is Monsell's (1985, Figure 5.4) finding that prior study in anticipation of a recognition test yielded facilitation for nonwords in his LDT. However, this facilitation was not a long-term effect, in that it occurred only when no items intervened between the nonword's study and test presentations.

(Manuscript received June 27, 1985;
revision accepted for publication June 12, 1988.)