Journal of Experimental Psychology:
Learning, Memory, and Cognition
1992, Vol. 18, No. 6, 1251-1269

0278-7393/92/$3.00

Direct Comparison of Two Implicit Memory Tests: Word Fragment and Word Stem Completion

Henry L. Roediger III
Rice University

Mary Susan Weldon
University of California, Santa Cruz

Michael L. Stadler
University of Missouri—Columbia

Gregory L. Riegler
University of St. Thomas

In 3 experiments, the implicit memory tests of word fragment and word stem completion showed comparable effects over several variables: Study of words produced more priming than did study of pictures, no levels-of-processing effect occurred for words, more priming was obtained from pictures when Ss imaged the pictures' names than when they rated them for pleasantness, and forgetting rates were generally similar for the tests. A different pattern of results for the first 3 variables occurred under explicit test conditions with the same word fragments or word stems as cues. We conclude that the 2 implicit tests are measuring a similar form of perceptual memory. Furthermore, we argue that both tests are truly implicit because they meet Schacter, Bowers, & Booker's (1989) retrieval intentionality criterion: Levels of processing of words have a powerful effect on explicit versions of the tests but no effect on implicit versions.

Implicit memory tests are those in which retention is indexed by transfer from previous experiences to performance on a task that (typically) seems unrelated to the previous experiences. In one form of these tests, subjects are exposed to material (often words or pictures) during a study phase, are given degraded stimuli (fragmented words or pictures) during a test phase, and are asked to name stimuli. Subjects show a benefit in performance in naming the degraded stimuli if they have recently studied the items, in relation to naming nonstudied items. This benefit is referred to as *priming*, and it indexes retention. Performance on such implicit tests is frequently dissociated from that on explicit tests requiring conscious recollection, such as recall and recognition, by a variety of subject variables and independent variables (Roediger, 1990; Schacter, 1987).

A variety of implicit tests has been used to investigate the relation between implicit and explicit forms of retention. The starting point for their development was in neuropsychological studies of memory; amnesic patients showed preserved retention on implicit tests but poor recollection when measured by explicit tests. The original assumption was that all implicit tests tapped the same kind of mnemonic ability, one underlain by a memory system that was spared in amnesic patients (Shimamura, 1986).

Roediger and Blaxton (1987b) suggested that implicit memory tests were not all alike and could themselves be dissociated. In particular, they suggested that some tests seemed to be based more on perceptual aspects of mental processing (data-driven tests), whereas others were based more on the meaning of events (conceptually driven tests), as first noted by Jacoby (1983). The distinction between perceptual and conceptual forms of implicit memory tests has received broad support from many recent investigations revealing that numerous variables dissociate these tests (Blaxton, 1989; Hamann, 1990; Rappold & Hashtroudi, 1991; E. R. Smith & Branscombe, 1988; Srinivas & Roediger, 1990; see Roediger, Srinivas, & Weldon, 1989, for a review). Evidence supporting this separation of a perceptual form and a conceptual form of priming also exists in the neuropsychological literature (Schacter, 1990; Tulving & Schacter, 1990), as do dissociations between other forms of implicit tests (e.g., Butters, Heindel, & Salmon, 1990).

Perceptual (or data-driven) implicit memory tests are those that challenge the perceptual system by presenting information in degraded forms, either fragmented items or items shown very rapidly. Such tests include word stem completion (ele______), word fragment completion (e_e__an_), word identification (presenting *elephant* rapidly at test), or picture fragment identification (naming a fragmented picture of an elephant). Conceptual implicit tests include generation of words from an associate (*tusk*), or from a category name (*animals*), or by answering a general knowledge question ("*What animal did Hannibal use to help him cross the Alps during his attack on Rome?*"). In the case of all implicit tests, subjects are simply told to solve the problem they are given as well as possible. In the case of the conceptual tests, there is no visual similarity between the test format and the study format; rather, the connection between study and test is based on meaning of the concepts. On the other hand, priming in the perceptual tests does depend in part on resemblance of the studied events to the test forms (see Roediger, Weldon, & Challis, 1989, for further specification of this distinction).

This research was completed while the authors were colleagues at Purdue University. Various aspects of research were supported by National Institutes of Health Grant R01 HD15054 awarded to J. H. Neely and Henry L. Roediger III and by Air Force Office of Scientific Research Grant 91-0253 awarded to Henry L. Roediger III.

The article benefited by comments from B. H. Challis, M. J. Guynn, J. H. Neely, K. B. McDermott, D. L. Schacter, M. C. Smith, and E. Tulving.

Correspondence concerning this article may be addressed to Henry L. Roediger III, Department of Psychology, Rice University, Houston, Texas 77251-1892. Electronic mail may be sent to roedige@ricevm1.rice.edu.

A working assumption in the framework suggested by Roediger and Blaxton (1987b) is that perceptual implicit memory tests should be similarly affected by manipulation of independent variables. For example, the implicit tests of perceptual identification, word fragment completion, and word stem completion all show strong modality effects, such that visual presentation of items produces more priming than does auditory presentation (see Jacoby & Dallas, 1981, Roediger & Blaxton, 1987a, and Graf, Shimamura, & Squire, 1985, respectively). An exception to this expectation of similar effects on perceptual tests, but one predicted by theory, is that they should produce different patterns of priming when the test forms differ, such as in the comparison of primed word fragment completion and primed picture fragment naming. Indeed, words produce more priming than do pictures in the former case, but pictures produce more priming than do words in the latter, as expected (Weldon & Roediger, 1987, Experiment 4).

Some evidence exists showing that not all putatively implicit verbal perceptual tests behave similarly, and the aim of the present research was to examine this claim. We concentrated on two of the most popular tests: word stem completion and word fragment completion. A priori, one might expect that these tests could differ in their effects. The typical word stem completion test (e.g., Graf & Mandler, 1984), in which subjects are given the first 3 letters of words, is marked by (a) each stem having 10 or more possible completions, (b) subjects being able to complete each item, and (c) responses being fast and the task being relatively easy. On the other hand, in word fragment completion (e.g., _o_o_ut for *coconut*; Tulving, Schacter, & Stark, 1982) (a) each fragment has only 1 or 2 possible completions, (b) subjects are unable to complete many test items, and (c) responses seem more effortful and typically occur more slowly. In addition, the materials typically used in word stem completion studies are high-frequency, short words, whereas those used in studies of word fragment completion are usually low-frequency, long words. Any or all of these differences could make these two measures—assumed to be equivalent by many researchers—differentially sensitive to experimental manipulations.

One difference often observed between word stem and word fragment completion is in the time course of forgetting: Primed word stem completion usually drops to baseline levels within 2 hr (e.g., Graf & Mandler, 1984), whereas primed word fragment completion persists for at least 1 week (Komatsu & Ohta, 1984; Roediger & Blaxton, 1987a; Tulving et al., 1982) and perhaps much longer (Sloman, Hayman, Ohta, Law, & Tulving, 1988). Graf and Mandler suggested that "word fragment completion may also involve deliberate retrieval whereas the stem completion test reflects automatic accessibility," because "[w]hen no completion immediately comes to mind, subjects may attempt to retrieve words they have recently seen" (p. 562). In short, word fragment completion may not be a pure implicit test, because it is contaminated by explicit retrieval strategies. Long-lasting priming on this test would therefore be an artifact.

A similar theme was sounded by Squire, Shimamura, and Graf (1987), who examined levels-of-processing effects in word stem and word fragment completion in both normal and amnesic subjects. They reported that amnesic patients did not show fully preserved priming on the word fragment completion test and that normal subjects (but not amnesic subjects) showed a slight levels-of-processing effect on the test. Both these findings in normal subjects differ from those reported with the typical word stem completion procedure, although slight levels-of-processing effects are usually also found in stem completion (see Challis & Brodbeck, 1992). Squire et al. (1987) suggested that word fragment completion tests are contaminated by a conscious recollection of the words during the test. They pointed out that no evidence had been reported of long-lasting priming on the fragment completion test with amnesic patients, who presumably could not use explicit retrieval strategies to accomplish the task.

More recently, Tulving, Hayman, and Macdonald (1991) have reported strong evidence of durable priming on a word fragment completion test in a densely amnesic patient, K.C. If anything, K.C.'s levels of priming on the standard word fragment completion test were higher than those of normal subjects, and his priming persisted for months. Of course, this finding still leaves the puzzle of why amnesic subjects (and normal subjects) show more durable priming on word fragment than on word stem completion tests, but Tulving et al.'s procedures involved more than the standard single presentation and test of material.

One possible solution to these differences between word stem and word fragment completion lies in an inherent confounding existing in all prior comparisons: word frequency (and hence familiarity and length). The studies previously described involved cross-experiment comparisons in which different materials were used for word stem and word fragment completion. As previously mentioned, materials used for word stem completion involve relatively short, high-frequency words, whereas those typically used for word fragment completion are long, low-frequency words. Much evidence exists showing that priming in the lexical decision task is greater for low-frequency than for high-frequency words (e.g., Scarborough, Cortese, & Scarborough, 1977), and the same seems to hold true for perceptual identification (Jacoby & Dallas, 1981) and for word fragment completion (MacLeod, 1989; but see Tenpenney & Shoben, 1992).

Because of this confounding, experiments are needed to directly compare word fragment and word stem completion on the same materials across a variety of experimental manipulations. To our knowledge, only two experiments have appeared in the literature directly comparing word fragment and word stem completion under both explicit and implicit test instructions. Nelson, Canas, Bajo, and Keelean (1987, Experiment 4) provided a direct comparison of word fragment completion to a variant of the word stem completion test, one in which word endings (e.g., ____oud for *cloud*) were given. The variables manipulated were the lexical set size of the target word (the number of words that would fit the cue) and the meaning set size (the number of associates generated to the target word). Lexical set size affected performance for both word fragment cues and word ending cues under both explicit and implicit test instructions. Nelson et al. concluded that processes involved in lexical access are critical for performance with both types of cues. However, the effect of meaning set size was different for the two types of cues; this variable did not affect performance with word fragments under either explicit or implicit instructional sets but did affect performance with word stems as cues with both types of instructions. It remains to be seen if Nelson et al.'s results would be replicated with the more typical word stem beginnings as cues, but the conclusion from their experiments is that word fragment completion is relatively immune to manipulations of meaning set size (see, too, Nelson, Keelean, & Negrao, 1989).

Weldon, Roediger, and Challis (1989, Experiment 1) gave subjects a mixed list of words and pictures under identical study conditions and then tested subjects under one of four conditions. Two groups of subjects were given explicit memory tests, with either word fragments or word stems that represented the studied items as cues. (Fragments or stems for the pictorial items referred to the pictures' names.) Two other groups were given the same cues, but they received instructions to say the first word to come to mind that fit the cue (implicit instructions). On the implicit tests, Weldon et al. reported greater priming from prior study of words than pictures for both word fragment and word stem completion. (The word fragment completion findings replicated results of Weldon and Roediger, 1987.) Although overall levels of performance were much higher on the explicit tests, as expected, the same pattern of results held: Words were better remembered than were pictures on cued-recall tests with both word fragments and word stems as cues.

Weldon et al. (1989) concluded that word fragments and word stems engaged similar (perceptual or lexical) retrieval processes, at least across the variables examined (study of pictures or words and test with explicit or implicit instructions). However, the limitations on this conclusion are great, because (a) it hinges on accepting the null hypothesis and (b) the number of variables is small. In addition, no dissociation was obtained between explicit and implicit versions of the task when test cues were held constant.

The reported experiments were designed to extend the conclusions arising from Weldon et al. (1989) to a wider range of conditions. In particular, we used a levels-of-processing manipulation to assess the extent to which explicit processes may contaminate the allegedly implicit word fragment completion test. Graf and Mandler (1984, Experiment 3) gave subjects word stems as cues following study of words under a typical levels-of-processing manipulation, with subjects' attention directed at either the words' sound or meaning during study. When subjects were given explicit retrieval instructions, Graf and Mandler found a large levels-of-processing effect, but when the same word stems were given with implicit instructions (i.e., complete the stems with the first words that come to mind), little or no levels-of-processing effect was obtained. Therefore, the magnitude of levels-of-processing effects under implicit test instructions on perceptually based tests can serve to calibrate how contaminated implicit tests are by explicit retrieval processes. Schacter, Bowers, and Booker (1989) have referred to this logic, in its general form, as the retrieval intentionality criterion for separating explicit and implicit tests: A different pattern of results is obtained when subjects are intending to retrieve (explicit tests) than when retrieval is an incidental by-product of performance (implicit tests).

Bowers and Schacter (1990) conducted further research with the word stem completion task and showed that little or no levels-of-processing effect occurred on the task when subjects were unaware of the relation between the study and test phases or when they were told before the test that some stems represented studied words but they should nonetheless respond with the first word that came to mind. However, for subjects who were nominally unaware of the relation between the study and test phases, but who caught on to the relation and presumably used word stems as recall cues, a levels-of-processing effect occurred. Bowers and Schacter's experiments provide further evidence that the levels-of-processing effect can be used as an index of whether an ostensibly implicit test is contaminated by explicit retrieval.

Experiment 1

Subjects studied a mixed list of pictures and words under conditions in which they made decisions about the appearance of the verbal form of the items or decisions about the items' meaning. At the test phase, four independent groups of subjects received word stems or word fragments under explicit or implicit test instructions.

Let us first consider predictions for studied words. We expected to replicate the Graf and Mandler (1984) findings with word stems as cues: A levels-of-processing effect should occur on the explicit form of the test but not on the implicit form of the test. The more interesting case involves word fragments as cues. If Graf and Mandler (1984) and Squire et al. (1987) are correct that presenting word fragments at the test phase necessarily creates explicit retrieval processes (even with implicit test instructions), then a levels-of-processing effect should be obtained with word fragments under both implicit and explicit instructions. On the other hand, if word fragment completion is an implicit, perceptually based test like word stem completion, then no levels-of-processing effect should occur on the word fragment completion task. According to the logic of the retrieval intentionality criterion, the word fragment completion task would be deemed a true implicit test.

The predictions for picture items are more complex. Of course, we predicted that more priming on the implicit tests should occur for words than for pictures, replicating the results of Weldon and Roediger (1987) and Weldon et al. (1989). However, we predicted that on the implicit versions of word stem and word fragment completion, a different kind of levels-of-processing effect should occur: Pictures encoded under the graphemic analysis condition should produce more priming than those encoded under the semantic (pleasantness rating) condition. The reasoning is as follows: When subjects studied pictures in the graphemic processing task, they were instructed to imagine the appearance of the word and to count the ascending and descending letters in the words. In doing so, subjects should create, albeit in imaginal form, percepts relevant to completing word fragments or word stems. If these tests are perceptually based, as claimed by Roediger (1990), then greater priming should occur after this graphemic processing than after the semantic processing (which did not involve imagining the appearance of the pictures' names). Some prior work gave us hope of obtaining this finding, because Donnelly (1988), Jacoby and Witherspoon (1982), and Roediger and Blaxton (1987a) found that when subjects were given words auditorily, but asked to spell the words or to imagine their visual forms, greater priming was found on word identification tasks than from uninstructed auditory presentations.

One final manipulation was used in Experiment 1: Half the subjects studied material under intentional learning conditions and half under incidental learning conditions. In many prior experiments, including our own, subjects studied material under intentional learning conditions and then received an implicit test such as word fragment completion under the guise that it was a filler task given prior to the memory test. However, because subjects expected to be tested on the material in the list, they might have realized that the implicit test involved studied list items and therefore used explicit strategies. We doubted this possibility but included both intentional and incidental learning conditions in the experiment to check it. Bowers and Schacter (1990), Greene (1986), and Neill, Beck, Bottalico, and Molloy (1990) all found no effect of this variable on primed word stem completion. The conditions of special interest involved performance on the implicit word fragment and word stem completion tests after incidental or intentional learning during study.

Method

Subjects and design. The subjects were 240 Purdue University undergraduates who earned credit in Introductory Psychology for participating. All were native English speakers with corrected or normal vision.

The design comprised a mixed factorial with five factors. The between-subjects factors were learning instructions (incidental or intentional), test instructions (implicit or explicit), and type of cue at test (word fragments or word stems). The within-subjects factors were the experimental history of the target (test) items (study as picture or word or nonstudied) and the level of processing (graphemic or semantic). Items were counterbalanced so that they served in all study and test conditions equally often across subjects.

Materials and list construction. Sixty target items (concepts) were selected from the Snodgrass and Vanderwart (1980) and Weldon and Roediger (1987) item sets. Two slides were prepared for each concept: one containing a black and white line drawing and the other containing the word typed in black lowercase letters in a font matching that of the fragments and stems used on the tests. Previous pilot work on a separate group of 40 subjects indicated that subjects provided the intended labels for the pictures 97% of the time with individual items ranging from 81% to 100%.

The items were divided into two sets of 30. Each set was further divided into three subsets that were rotated through the picture, word, and nonstudied conditions within both the graphemic and semantic processing conditions. Half the subjects studied one set with the graphemic processing task and the other set with the semantic processing task; the assignment was reversed for the other subjects. Level of processing was blocked so that each subject received two separate lists: one with the graphemic task and one with the semantic task. The order of blocks was counterbalanced so that half the subjects performed the graphemic task first and half performed the semantic task first. Each study list contained 10 pictures, 10 words, and 6 buffer items (3 at the beginning and 3 at the end). The 26 words and pictures were randomly ordered in the list with the exception that no more than 3 words or pictures appeared sequentially.

The fragment and stem completion tasks contained the 60 target items and 32 fillers. Thus, fewer than half the items on the tests were previously studied (i.e., 40 studied targets, 20 nonstudied targets, and 32 fillers). Each fragment was intended to have a unique solution, although a few turned out to have two solutions; tests on separate groups of subjects indicated that the average baseline rate of completion was 33% with individual items ranging from 17% to 52%. Each stem had at least four alternative completions with different root morphemes in *Webster's New International Dictionary* (1934). Most word stems had many more completions. The first 10 test items were fillers, and the remaining fillers and fragments (or stems) were randomly arranged. Target items and fragments are presented in Appendix A.

Procedure. Subjects were tested in small groups of 1 to 5. They were told that the purpose of the session was to help prepare materials for future research projects and that they would perform several different tasks. In addition, subjects in the intentional learning conditions were told that they would receive a memory test for the pictures and words seen in the first 2 tasks (the semantic and graphemic processing conditions), but the nature of the test was not specified. Subjects in the incidental learning groups were not told that they would receive a memory test.

Subjects then received instructions for the first processing task. For the graphemic task, subjects were told that they would see a series of pictures and words on slides. For each item, they were to count the total number of letters in the word, or in the name of the picture, that had ascending letters (e.g., b, d, f, h, k, l, and t) and descending letters (e.g., g, j, p, q, and y) and to circle a number between 0 and 7 on their rating sheet to record their response. Subjects were shown examples of words with ascenders and descenders. For the pictures, they were advised to form a mental image of the word or to spell it and count the ascenders and descenders on their fingers.[1] Four practice items (two pictures and two words) were presented and scored to be certain that everyone understood the task. Subjects in the intentional learning group were reminded to study the items for the memory test. The study list was then presented at a rate of 10 s per item with subjects responding to each item on their rating sheet.

In the semantic task, subjects were told that they would rate a series of pictures and words for pleasantness. They were told that as they saw each item on a slide, they should think about the real-world object and then rate it for pleasantness on a scale of 0 (*extremely unpleasant*) to 7 (*extremely pleasant*) by circling a number on their rating sheets. They were given four practice items (two pictures and two words). Subjects in the intentional learning group were reminded to study the items for the memory test. The study list was presented at a rate of 10 s per item with subjects performing the rating task. Note that the overt features of the study tasks (rating items on a 0 to 7 scale) were the same for the two orienting tasks, but the type of judgment differed, thus instantiating the levels-of-processing manipulation.

After performing the semantic and graphemic processing tasks, subjects received two filler tasks. On the first task they wrote as many names of the U.S. presidents as they could remember in 5 min. Next,

[1] Referring to performance of this task with pictures as graphemic encoding is something of a misnomer, because subjects actually had to covertly name the picture, imagine the name typed, and then count ascenders and descenders (as with the words). Even though performing the graphemic task clearly involves different processes for words and pictures, we use the same term in the tables and figures to refer to both. The actual overt behavior (judging the number of ascenders and descenders) was the same in the two cases.

they completed the short form of the Need for Cognition survey (Cacioppo & Petty, 1982) for 3 min. These filler tasks were intended to help disguise the nature of the implicit tests by encouraging subjects to perceive all of them as a series of unrelated tasks.

Subjects then received the appropriate test. In all cases, word stems or word fragments were presented at a 12-s rate. In the implicit test conditions, subjects were not told that some of the test items were related to the previously studied pictures and words, but they simply were told that they were going to solve word puzzles. Subjects in the word fragment completion group were told that they would see word fragments with missing letters and that they should write the first word that came to mind that successfully completed the fragment. Subjects in the word stem completion group were instructed to write the first word that came to mind that began with the three letters provided. Both groups were shown an example of the appropriate type of test item. They were told that they would have 12 s to solve each item and that they should neither work ahead nor return to any items. They were given a cover sheet to conceal the upcoming items. The task was timed with an audio tape on which the word *next* was spoken every 12 s.

Subjects in the explicit test condition were told that they were going to receive a memory test for the pictures and words they had seen on the slides. They were told that some of the fragments (or stems) were clues to help them remember the items and that they should look at each clue and try to remember the name of a picture or word from the slides that would complete the fragment (or stem). They were also told that many of the fragments (stems) did not refer to a previously studied item and that they should leave these lines blank. They were told not to guess and to be relatively certain that they had studied the item before they wrote down its name. As with the implicit test group, subjects were given 12 s to retrieve each item, and they were told neither to work ahead nor to return to any fragments (stems). They also were provided with a cover sheet. The test sheets of word fragments and word stems were the same for both explicit and implicit conditions; only the test instructions differed.

Results

The results are complex, owing to the 5-factor design used. We first present overall analyses and then consider specific comparisons of interest. The overall results are presented in Tables 1 and 2. Table 1 shows performance on the implicit memory tests for words and pictures under both levels-of-processing conditions and both incidental and intentional learning conditions and for both word fragments and word stems as cues. The top half of Table 1 shows total proportion completed, whereas the bottom half shows priming scores (i.e., the base rate for completing fragments or stems in the different groups was subtracted out). Examination of the data at the bottom of Table 1 shows that (a) more priming occurred for words than for pictures on both tests, (b) no levels-of-processing effect occurred for either test following study of words, (c) greater priming occurred for pictures when subjects imaged their names than when they judged pleasantness, and (d) there was no effect of intention to learn the materials. Priming may be higher on word fragment completion than on word stem completion because on average more letters are given in fragments than in stems.

Table 2 shows the basic data for the explicit test condition, uncorrected for guessing. The intrusion rates (the proportion of nonstudied items mistakenly recalled) are shown on the right and are generally fairly low; they are much lower than the comparable nonstudied rates on the implicit tests. The following trends are apparent in the data: (a) a sizeable levels-of-processing effect following study of words; (b) either no effect or a tendency toward a reverse effect in the study of pictures; (c) much higher performance on the explicit test (Table 2) than on the implicit test (Table 1), even if intrusions

Table 1
Proportions of Test Fragments and Stems Completed Under Implicit Retrieval Instructions in Experiment 1

	Study condition				
	Words		Pictures		
Cues/study	Graphemic	Semantic	Graphemic	Semantic	Nonstudied
		Total completed			
Word fragment					
Incidental	.51	.48	.41	.29	.28
Intentional	.52	.50	.39	.30	.27
M	.52	.49	.40	.30	.28
Word stem					
Incidental	.31	.28	.20	.18	.15
Intentional	.31	.35	.25	.19	.17
M	.31	.32	.22	.18	.16
		Priming			
Word fragment					
Incidental	.23	.20	.13	.01	—
Intentional	.25	.23	.12	.03	—
M	.24	.22	.13	.02	—
Word stem					
Incidental	.16	.13	.05	.03	—
Intentional	.14	.18	.08	.02	—
M	.15	.16	.06	.02	—

Note. Dashes indicate that priming scores, by definition, do not exist for nonstudied items.

Table 2
Proportions of Test Fragments and Stems Completed Under Explicit Cued-Recall Instructions in Experiment 1

	Study condition				
	Words		Pictures		
Cues/study	Graphemic	Semantic	Graphemic	Semantic	Nonstudied
Word fragment					
Incidental	.31	.51	.42	.36	.09
Intentional	.26	.49	.36	.35	.04
M	.28	.50	.39	.35	.06
Word stem					
Incidental	.16	.59	.51	.43	.02
Intentional	.22	.56	.48	.46	.05
M	.19	.58	.50	.45	.04

are subtracted from the former data; and (d) little or no effect of intentionality of learning.

Basic statistical analyses were computed on difference scores to take baseline rates out of the picture. These were priming scores for the implicit tests (studied − nonstudied completion rates) and scores corrected for guessing for the explicit tests (correct − intrusions). The use of difference scores is controversial, but this procedure seemed preferable to including nonstudied base rates in the analyses of variance (ANOVA), because numerous artifactual interactions are produced by doing this.

An initial finding from the 5-way ANOVA helps to simplify presentation of the remainder of the results and the supporting statistics. There was no significant main effect of type of study (intentional or incidental) and no interaction involving this variable that reached conventional levels of statistical significance. After collapsing the data over all other conditions in the experiment (and correcting for base rate), we found the proportion of responses produced after incidental learning conditions to be .23 and after intentional conditions to be .24. Therefore, all other statistical analyses were conducted using a 4-way ANOVA, collapsing over this variable. However, we have chosen to simplify matters further by putting all inferential statistics in Appendix B and reporting only the reliable trends in this section.

The rationale for omitting such statistics is as follows: Inclusion of the between-subjects factor of study instructions (intentional or incidental), which had no effect, obviates the need for reporting inferential statistics to the extent that the same results were obtained after both study instructions. All findings in the experiment were directly replicated, thus eliminating the need to perform inferential statistics to know if claimed effects are indeed reliable. (The 120 subjects who studied under intentional conditions can be considered a direct replication of the experiment in which the 120 subjects studied under incidental conditions, and vice versa.) Further evidence on the reliability of the observations can be seen in the high correlations between intentional and incidental learning conditions across the eight other conditions represented in both explicit and implicit tasks (Pictures/Words × Graphemic/Semantic Processing × Word Stems/Word Fragments). For the implicit tests, the correlation between intentional and incidental study conditions was .95, whereas it was .96 for the explicit conditions. Skeptics of this logic may consult F ratios in Appendix B and may wish to consider a similar case reported by Tulving (1983, pp. 204–205). We now bring out the main features of interest in the results by considering first the implicit and then the explicit test conditions.

Implicit tests. Two groups of subjects were given implicit memory tests after the various encoding conditions (study of words or pictures under graphemic or semantic orienting tasks). The composite results (combined over incidental and intentional study conditions) are presented graphically in Figure 1 for ease of comprehension. For both tasks, there is much greater priming from words than from pictures when these items were presented during study under semantic encoding conditions, replicating past work (Weldon & Roediger, 1987; Weldon et al., 1989). A second finding of interest is that no levels-of-processing effect occurred for word stimuli. The word stem completion results replicate those of others who have found little or no levels-of-processing effect on word stem completion (e.g., Graf, Mandler, & Haden, 1982; Graf & Mandler, 1984) or on other perceptual implicit tests (Jacoby & Dallas, 1981). The finding of no levels-of-processing effect on the word fragment completion test is new. Coupled with the finding of a powerful levels-of-processing effect on explicit word fragment cued recall described later, the absence of a levels-of-processing effect implies that word fragment completion is truly an implicit memory test, contrary to the claims of Graf and Mandler (1984) and Squire et al. (1987), among others.

A third interesting finding revealed in Figure 1 is that pictures graphemically encoded (see Footnote 1) produced more priming on both the word fragment and word stem completion tests than did pictures studied under semantic coding conditions. Although such a reverse levels-of-processing effect is counterintuitive from most perspectives, it is perfectly consistent with the transfer-appropriate processing viewpoint (Roediger, 1990). Indeed, a similar result in explicit memory tests helped initiate this framework (Morris, Bransford, & Franks, 1977). When subjects studied pictures under the graphemic orienting task, they were asked either to imagine what the name of the picture looked like or to spell it out mentally while counting ascenders or descenders. Thus, subjects had to visualize the word to perform the orienting task,

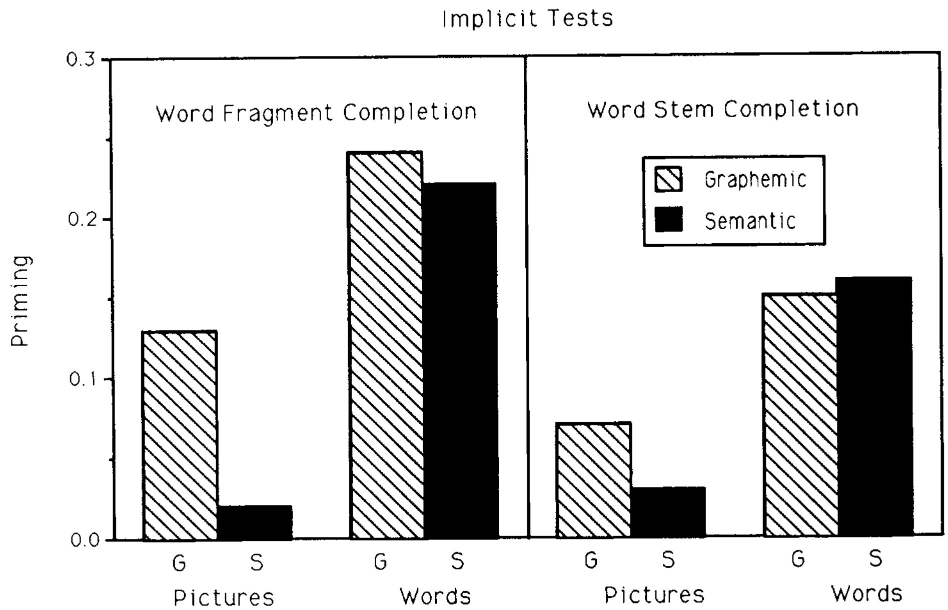

Figure 1. Priming on the implicit memory tests in Experiment 1. (Bars represent the benefit in completing fragments or stems above the nonstudied base rates.)

and their images aided performance on the word fragment and word stem completion tasks. These data agree with those of Donnelly (1988), Jacoby and Witherspoon (1982), and Roediger and Blaxton (1987a), described earlier, and bolster the claim that these tests rely on perceptual operations directed at reading words or on lexical processing (see Weldon, 1991). Forming images of the referents of words (rather than the visual appearance of words) has little or no effect on primed word fragment completion under most circumstances (Blaxton, 1989, Experiment 3).

Explicit tests. Two other groups of subjects received explicit cued-recall tests after the various study conditions with either word fragments or word stems as recall cues. Results are shown in Figure 2 (based on the data in Table 2). The most important result to glean from these data is that a strong levels-of-processing effect emerged for words on both explicit

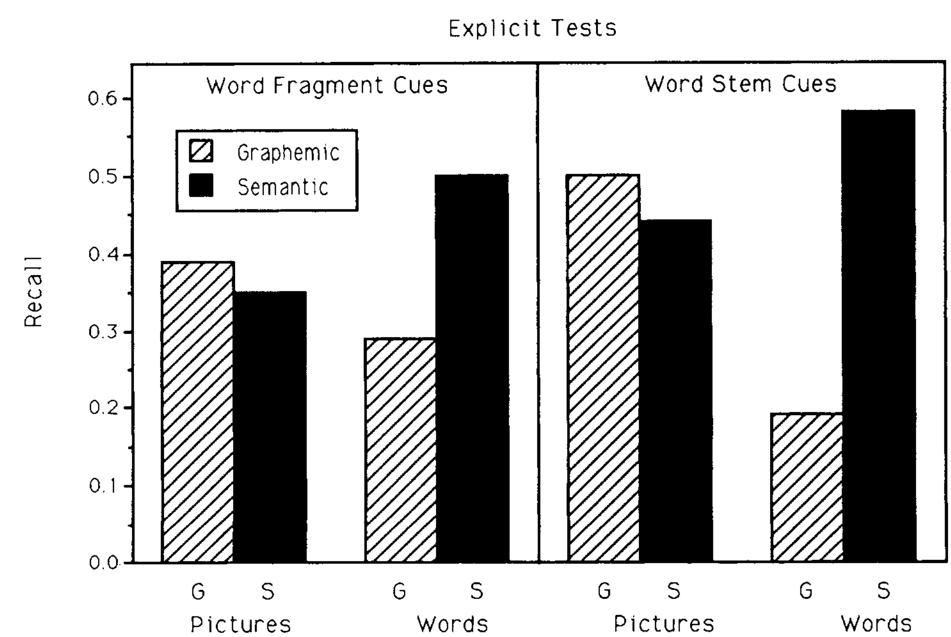

Figure 2. Recall on the explicit memory tests in Experiment 1. (Bars represent cued recall following the various study conditions with guessing rates subtracted out.)

tests. Like Graf and Mandler (1984, Experiment 3), we find a strong levels-of-processing effect on cued-recall of words with word stems as cues, coupled with a lack of effect on an implicit word stem completion test (see Figure 1). The new feature in the present data is that exactly the same pattern occurs with word fragments as cues: a strong levels-of-processing effect on an explicit version of the test (see Figure 2), but no effect on an implicit version (see Figure 1), with all other aspects of study and test held constant. Thus, both implicit tests meet Schacter et al.'s (1989) retrieval intentionality criterion.

Two other features of the data in Figure 2 are worth noting. First, under semantic encoding conditions, words produced better recall than did pictures with both word fragments and word stems as retrieval cues. The word fragment cued-recall data are consistent with those of Weldon et al. (1989) and the better recall of words may be attributed to the data-limited nature of the fragment cues, which impose a strong perceptual constraint on responding (see Weldon et al., 1989). The second finding is new and more interesting: Under graphemic encoding conditions, pictures are better remembered than words. This finding can be understood in the context of two factors. As just mentioned, the data-limited nature of fragments and stems support a perceptual base of retrieval in these tasks, and the graphemic orienting task forced subjects to imagine the word form during study, which in turn aids recovery of the word when the fragment or stem is given as the recall cue. (This part of the interpretation is the same as that given for the implicit test results.) However, Roediger, Weldon, and Challis (1989, p. 34) suggested that explicit retrieval instructions encourage meaning-based retrieval processes in recalling events. This assumption seems demanded by the finding (here and elsewhere) that the levels-of-processing manipulation has large effects on explicit recall with word fragments and word stems as cues. Because meaning of pictures seems to be more automatically coded than meaning of words (e.g., M. C. Smith & Magee, 1980), we may assume that more meaningful processing occurs with pictures than with words even under graphemic encoding conditions. In summary, the picture superiority effect following the graphemic orienting task with both explicit tests can be attributed to two factors: Imaging the word benefits later recovery of the study episode from the verbal cue (in relation to semantic encoding of the pictures), and the greater meaning extracted from pictures than words also supports retrieval on an explicit test that is assumed to rely partly on meaning-based retrieval. These two factors can be interpreted in the context of a generation/recognition model of performance for these tasks as suggested by Jacoby and Hollingshead (1990), in which the generation phase involves recovering a word from its fragmented form (and is affected by perceptual factors) and the recognition phase consists of conceptual processes invoked in deciding whether generated words appeared in the study list. Further discussion of the implications of these results for the generation/recognition model is deferred until the General Discussion section.

Experiment 2

The results of Experiment 1, and those of Weldon et al. (1989), indicate that word fragments and word stems produce quite similar patterns of performance both on implicit and explicit memory tests. The similarity on implicit tests is of most interest. The observed similarities include (a) more priming from words than from pictures, (b) no levels-of-processing effect following study of words, and (c) greater priming following study of pictures under word imaging and graphemic encoding than semantic encoding conditions. All three outcomes strongly support the conclusion that tests under implicit instructions with word stems or word fragments as cues are truly implicit, with a perceptual (lexical) basis, and are little affected by encoded meaning.

Although the two forms of implicit test are quite similar across the three variables of Experiment 1, we have not yet examined the variable that has led to most speculation about differences among tests: retention interval. Most prior work has shown that primed word stem completion decays rapidly and often disappears within 2 hr (e.g., Graf, Squire, & Mandler, 1984). On the other hand, primed word fragment completion certainly lasts 1 week (Komatsu & Ohta, 1984; Roediger & Blaxton, 1987a; Tulving et al., 1982) and probably even 1 year, at least after repeated exposures of material (Tulving et al., 1991). However, description of the effects of forgetting in priming tasks is in disarray because some reports have shown significant priming after 2 hrs or longer in the word stem completion task (e.g., Squire et al., 1987, Experiment 2; Warrington & Weiskrantz, 1978) and others have shown rather rapid declines even over short intervals in primed word fragment completion (Sloman et al., 1988).

The major difficulty in interpreting prior research comparing forgetting functions on implicit memory tasks is that all previous comparisons have been across experiments. As mentioned in the introduction, virtually all comparisons are confounded with word frequency, familiarity, and length. Researchers investigating word stem completion usually use relatively short, high-frequency words, whereas those investigating word fragment completion have used long, low-frequency words. Differences that have been attributed to the nature of the implicit test may instead be due to the frequency, familiarity, or length of the different types of words used in the research.

Experiments 2 and 3 were designed to gain more information on the decay rates of priming in implicit memory tests with other factors held constant. Subjects in Experiment 2 received a long list of words under incidental learning conditions and immediately took a test. Half the subjects returned after 90 min to take a second test; for the other half, the second test occurred after 48 hr. The tests used were word fragment completion, word stem completion, and anagram solution. The last test involves having subjects construct words from letter strings (e.g., *yhobodo* to *boyhood*) and priming is revealed by subjects solving more items if they have been recently presented than if not (e.g., Schoen, Ciofalo, & Rudow, 1989). We included this test in Experiment 2 because it was under active investigation in other projects in our laboratory at the time this experiment was being planned (Srinivas & Roediger, 1990).

Method

Design and subjects. The subjects were 120 Purdue University students in Introductory Psychology who received credit toward an experimental participation requirement. After being presented with a list of words, they were given one of three implicit memory tests both immediately after presentation and after a longer delay. The three tests were word fragment completion, word stem completion, and anagram solution. All subjects were tested immediately after presentation and then retested. On the retest, half of the subjects were tested after 90 min and the other half after 48 hr. Thus, the design comprised a 3 (type of test) × 3 (delay) factorial, although for analytic purposes we treated the design as a 3 × 2, because all subjects did not participate in both delay conditions.

Materials. A pool of 136 words was selected from various sources, including Blaxton (1989), Durgunoglu and Roediger (1987), Tulving et al. (1982), and Weldon and Roediger (1987). All words have seven or fewer letters. The entire set of words is displayed in Appendix C, along with the word fragments and anagrams used on two of the implicit tests. We did not require that all fragments have single solutions to be included, although in fact 99 (71%) did. All but 5 had fewer than 3 completions. Word stems had at least 5 completions according to the *Random House Dictionary* (1984), and 90% had more than 10 completions. (The mean number of completions was greater than 10 for the entire set.) Anagrams were formed by randomly arranging letters from the words.

The 136 items were randomly divided into four groups of 34. The item sets were systematically rotated through the four conditions on each implicit test (studied or nonstudied and immediate test or delayed test). This procedure served to balance specific items over all conditions of the experiment.

Procedure. Subjects were tested in groups of 1 to 10 people and were told at the outset that the experimenters were collecting norms about the characteristics of words for future psycholinguistic research. Subjects were shown 68 words presented by a Kodak Carousel slide projector at a 7-s rate and were instructed to judge each word for its pleasantness by circling a number from 1 (*very unpleasant*) to 7 (*very pleasant*) on their response sheets.

After presentation of the list, subjects were given a distractor task (recalling as many U.S. states and capitals as possible) for 6 min. After this distractor task, subjects were instructed to perform one of the three implicit memory tasks (word stem completion, word fragment completion, or anagram solution). Instructions for word fragment and word stem completion were similar to those used for the implicit test conditions in Experiment 1. Subjects receiving the anagram solution test were also not told about any relation between test items and the words given in the first phase of the experiment; they were told that they would see words with their letters randomly arranged and that their task was to unscramble the letters, if possible, to form a word in the allotted time.

Subjects in all test conditions were paced by a prerecorded signal ("next") played on a tape recorder such that they were allowed 12 s for each test item. A cover sheet was used to hide upcoming test items, and subjects were told to work only on the current test item and not to return to any items that had not been solved in the allotted time. Each test included 68 items, half of which had been studied, presented in a random order.

Forty subjects took the initial test in each of the three conditions. Half of these subjects returned 2 hr later for a second test similar to the first test, whereas the other half of the subjects returned 48 hr later for the second test. The delayed test was always like the first in terms of instructions and in that 68 items (half studied and half nonstudied) were included, but different items were tested on the two occasions. After their second test, subjects were debriefed.

Results and Discussion

The results of Experiment 2 are shown in Table 3, in which the proportion of items completed or solved is shown for all three implicit memory tasks at three delays. The data reveal significant priming at all delays in the experiment, even the .05 priming effect in the 48-hr delay condition in the anagram solution test, $t(38) = 2.04$. Both word fragment and word stem completion tests revealed significant priming effects after the 48-hr delay; indeed, if anything, there was greater loss over time in the word fragment completion task than in the word stem completion task, although this conclusion is compromised by the fact that the baselines and amounts of priming differed in the two tasks. At any rate, on the basis of these data, no great differences appear to exist in losses of priming over time in the word fragment and word stem completion tasks when other variables are not confounded with the tests.

Table 3
Proportions of Test Items Completed in Experiment 2

		Retention interval (h)		
Task	Condition	0	2	48
Word fragment completion	Studied	.52	.50	.50
	Nonstudied	.29	.30	.34
	Priming	.23	.20	.16
Word stem completion	Studied	.31	.30	.26
	Nonstudied	.17	.16	.15
	Priming	.14	.14	.11
Anagram solution	Studied	.41	.40	.38
	Nonstudied	.26	.25	.33
	Priming	.15	.15	.05

We conducted three separate ANOVAs on the data in Table 3 because of the unusual design. Recall that all subjects taking each test were tested in the 0-hr retention interval, but only half the subjects were tested at each delay (on different items than in the immediate test). Thus, the comparison between the 2- and 48-hr delays is between-subjects. We conducted an initial 2 (retention interval: 0 or 2 hr) × 3 (implicit memory test) ANOVA on the priming scores for the half of the subjects who did not receive the 48-hr test. This analysis produced a main effect of test, $F(2, 57) = 3.46$, $MS_e = 18.67$, indicating that more priming occurred in the word fragment completion test than in the other two implicit tests, but there was no main effect of delay, $F(1, 57) < 1$, $MS_e = 14.26$. Thus, no loss of priming occurred over a 2-hr period. However, it should be borne in mind that the first test at the 0-hr retention interval actually occurred 7–8 min after the list was presented, due to the filler task and instructions. In addition, the time between study and test of each individual item was longer because of the intervening study and test of other items. Thus, the amount of priming at the 0-hr retention interval inevitably reflects some loss of priming, too (see Sloman et al., 1988). Still, it is interesting that little further loss occurred over an additional 2 hr.

A second 2 (retention interval: 0 or 48 hr) × 3 (type of test) ANOVA for the half of the subjects that did not take the 2-hr test indicated a main effect of type of test, $F(2, 57) = 4.78$, $MS_e = 17.79$, and of retention interval, $F(1, 57) = 5.31$, $MS_e = 19.34$, but no interaction between these factors, $F(2, 57) = 1.98$, $MS_e = 19.34$. A significant drop in priming occurred between the initial test and the one occurring after 48 hr. We confirmed this loss of priming in a third 2 (2-hr vs. 48-hr delay) × 3 (type of test) ANOVA that compared performance for subjects taking the tests between 2 and 48 hr. Once again, there was a significant main effect of task, $F(2, 114) = 4.04$, $MS_e = 76.98$, and of delay, $F(1, 114) = 5.30$, $MS_e = 100.83$, but no significant interaction between these factors.

In summary, there was no statistically significant loss in priming on three implicit tests between an initial test given soon after study of a list and a test given 2 hr later. However, a significant loss of priming did occur between 2 and 48 hr. Furthermore, there was no dramatic difference between word fragment and word stem completion over the retention intervals used in Experiment 2. As in Experiment 1, these tasks behaved in a similar fashion as a function of an independent variable, but in this case (unlike those in Experiment 1), the variable is one claimed in past work to dissociate these tests. Experiment 3 was an attempt to confirm and extend this null result.

Experiment 3

The results of Experiment 2 showed that priming existed on both word fragment and word stem completions after 48 hr; prior research has shown that priming on word stem completion usually dropped to baseline levels in 2 hr. In Experiment 3, we sought to replicate the findings of Experiment 2—especially the significant priming in word stem completion after 48 hr—and to extend the retention interval to 1 week. In addition, in an attempt to reconcile the findings of Experiment 2 with the prior literature, we varied word frequency. Our hypothesis was that there would be less priming and that priming would decay to baseline levels of performance more rapidly for high-frequency than for low-frequency words. If so, this pattern would help explain why previous researchers obtained disparate effects (across experiments) when examining word fragment and word stem completion: Researchers studying word stems typically used short, high-frequency words as materials, whereas those studying word fragment completion used long, low-frequency words (e.g., Squire et al., 1987).

Prior research examining the effects of word frequency on priming has generally shown that low-frequency words produce more priming than do high-frequency words in the lexical decision task (Scarborough et al., 1977), in perceptual identification (Jacoby & Dallas, 1981), and in word fragment completion (MacLeod, 1989). However, Tenpenney and Shoben (1992) have recently presented some evidence for greater priming from high- than from low-frequency words in word fragment completion; therefore the matter is not settled.

Subjects in Experiment 3 were exposed to a long list of words and then received either an implicit word fragment completion test or word stem completion test shortly after study. Different subgroups of subjects then returned to take another test of the same type (on different items) after 90 min, 48 hr, or 1 week. Half the studied and tested words were high-frequency, and half were low-frequency.

Method

Subjects and design. The 144 subjects were Purdue University undergraduates who participated either as part of a course requirement or for pay. Half the subjects were tested on an implicit word fragment completion test and the other half received a word stem completion test under comparable conditions. All subjects in both test conditions received an initial test soon after seeing a long list of words. One third of the subjects in each test condition returned after 1.5 hr for a second test, one third returned after 48 hr, and one third returned after 168 hr. Half the items on each test were studied, and half were nonstudied; of each type, half were high-frequency words, and half were low-frequency words.

Materials. A set of 128 words was selected from the norms of Paivio, Yuille, and Madigan (1968) and Friendly, Franklin, Hoffman, and Rubin (1982); half were low-frequency and half were high-frequency words. (Sixteen other words served as filler items given during study.) The 64 low-frequency words had frequency counts under 10 per million in the Kucera and Francis (1967) norms, whereas the high-frequency words had counts of over 50 per million. The mean frequency values for the item sets were 3.7 per million and 118 per million. In addition, the words in the two sets were equated on word length and imagery ratings. High-frequency words averaged 7.3 letters per word and had imagery ratings of 4.4 (on the usual 7-point scale; 7 = *high imageability*); the comparable values for low-frequency words were 7.3 letters per word and an imagery value of 4.7. We should note that most prior experiments in which word frequency was manipulated did not equate items on these other characteristics.

The set of 128 target words and their corresponding fragmented forms are given in Appendix D. The fragments were not designed to have unique solutions, although 66 did. Of the remainder, 31 had two solutions, 20 had three to five solutions, and 11 had more than five solutions. The first 3 letters of each word were used as the stems. All stems could be completed by at least five words with different root morphemes, and the mean number of possible completions was greater than 10.

The 128 items were divided into four groups of 32, each with 16 high-frequency and 16 low-frequency words. Mean word length and imagery value were equated across the four groups of items. These item groups were rotated through the four study and delay conditions (i.e., studied and nonstudied items on the immediate test and studied and nonstudied items on the delayed test). Thus, across all subjects, each item was tested an equal number of times as a studied and nonstudied item in all delay conditions for both word fragment and word stem completion tests.

Procedure. The procedure in Experiment 3 was similar to that of the previous two experiments. Briefly, subjects examined a list of 80 words at a 7-s rate (64 target items and 16 fillers), judging each for its pleasantness on a 1–7 scale. After the same 6-min distractor task used in Experiment 2, all subjects received either a word fragment or word stem completion test under implicit instructions (i.e., say the first word that comes to mind that completes the stem or fits the fragment). Initially, 72 subjects took each type of test; for each delay condition, 24 subjects returned 1.5, 48, or 168 hr (1 week) later for the second test. Each test contained 64 fragments or stems, half studied and half nonstudied and half high-frequency and half low-frequency words. Subjects were given 12 s per fragment or stem on the test.

Results and Discussion

The basic data from Experiment 3 are presented in Table 4, which shows completion rates for studied and nonstudied items, as well as priming, for the word fragment completion test (top) and the word stem completion test (bottom). In each case there is evidence for a gradual decay in priming over time, but significant priming occurred for both tests at all delays (the smallest t was 1.98 for high-frequency words at the longest delay in word stem completion).

Table 4
Proportions of Test Items Completed in Experiment 3

Word frequency	Condition	Retention interval (h) 0	1.5	48	168	*M*
	Word fragment completion					
High	Studied	.38	.35	.40	.35	.37
	Nonstudied	.29	.22	.32	.27	.28
	Priming	.09	.13	.08	.08	.09
Low	Studied	.41	.39	.38	.30	.37
	Nonstudied	.22	.23	.24	.20	.22
	Priming	.19	.16	.14	.10	.15
	Word stem completion					
High	Studied	.38	.32	.32	.28	.32
	Nonstudied	.25	.22	.26	.23	.24
	Priming	.13	.10	.06	.05	.08
Low	Studied	.20	.22	.18	.12	.18
	Nonstudied	.08	.07	.07	.07	.07
	Priming	.12	.15	.11	.05	.11

We conducted two ANOVAs on the data in Table 4. First, we used a 2 (word frequency: low or high) × 2 (type of test: word stem or word fragment completion) × 2 (time of test: immediate or delayed) analysis. This analysis collapsed performance at all delay intervals, so that delay and word frequency were within-subjects factors and type of test was a between-subjects factor. There was a significant effect of word frequency, $F(1, 142) = 12.79$, $MS_e = 4.30$, with low-frequency words leading to greater overall priming (.14) than high-frequency words (.10). There was no significant effect of type of test, $F(1, 142) = 1.84$, $MS_e = 8.72$. The main effect of delay was significant, revealing decay over time, $F(1, 142) = 7.35$, $MS_e = 5.75$. The Word Frequency × Test interaction was also significant, $F(1, 142) = 6.01$, $MS_e = 4.30$, as was the third-order interaction among factors, $F(1, 142) = 5.61$, $MS_e = 4.61$. Briefly, word frequency had a greater impact on word fragment completion than on word stem completion, and priming declined over 1 week in all conditions except in word fragment completion with high-frequency words.

A second 2 (word frequency: low or high) × 3 (delay: 1.5, 48, or 168 hr) × 2 (test: word stem or word fragment) ANOVA was conducted without data from the immediate test. Word frequency was a within-subjects factor, whereas test delay was now a between-subjects factor. Again, there was a reliable main effect of frequency with greater priming from low-frequency words (.12) than from high-frequency words (.08), $F(1, 138) = 4.32$, $MS_e = 4.40$. The effect of delay was again significant, with priming declining from .13 to .10 to .07 from 1.5 to 48 to 168 hr, $F(2, 138) = 3.38$, $MS_e = 6.83$. The effect of test was not significant, $F(1, 138) = 2.35$, $MS_e = 6.83$, although priming was slightly greater in the word fragment completion test than in the word stem completion test, .11 to .08. No interactions between factors approached significance; all $Fs < 1$.

The results of Experiment 3 generally confirm those of Experiment 2. When the same materials are used for both types of test, there is little difference in loss of priming between word fragment and word stem completion. The significant priming seen in Experiment 2 after 48 hr was replicated and extended to 1 week, even in word stem completion. More priming occurred from low-frequency than from high-frequency words, although the magnitude of the frequency effect was not large numerically.

In some ways, the similarities in priming (and in the loss of priming) between the two tests was surprising, especially for low-frequency words. The nonstudied base rates for low-frequency words in the word stem completion test were much lower than for the same words in the fragment completion test. Because word stems afford the possibility of many completions, subjects usually think of a higher frequency response in the nonstudied condition. (Nonstudied base rates in the other three conditions—high-frequency items in both tests and low-frequency items in the word fragment completion test—were roughly comparable.) Despite the low base rate in this one condition, priming was about the same as in the other conditions.

General Discussion

The main findings from the three experiments were as follows

1. Following study of words under different orienting tasks, explicit cued-recall tests with either word fragments or word stems as cues showed large levels-of-processing effects, but on implicit versions of these tests levels-of-processing effects were completely absent.
2. Following study of pictures, there was greater priming (on implicit tests) or recall (on explicit tests) after the word imaging orienting task than after the pleasantness rating task.
3. Intentional versus incidental learning during study produced no main effect or interactions in Experiment 1.
4. In the two experiments examining possible differences in loss of priming over time between word fragment and word stem completion tests, little evidence of a differential effect was found. We found reliable priming on both word stem and word fragment completion tests after 1.5, 2, 48, and 168 hr. Priming was somewhat greater for low- than for high-frequency words on both tests. Although direct comparison of the two measures is hazardous because of differential base rates and other differences between tests, no obvious differences emerged in the results between the fragment completion and stem completion tests. We now discuss implications of the results for four issues in implicit memory: test differences, evaluation of tests, possible levels-of-processing effects, and information loss (or forgetting) on implicit tests. Finally, we discuss implications for Jacoby and Hollingshead's (1990) generation/recognition theory.

Comparison Between Word Fragment and Word Stem Completion

Graf and Mandler (1984) and Squire et al. (1987) have argued that word stem completion represents a purer measure of implicit memory than does word fragment completion, because responding is quicker and more automatic in the former case. They noted that priming on the word stem completion tests does not survive 2 hr in most experiments, whereas priming on the word fragment completion test lasts much longer. Previous comparisons between word fragment and word stem completion have usually been conducted between experiments, however, with numerous confoundings in materials and procedures (recall that Weldon et al., 1989, directly compared these measures, but they found no difference between tests in the variables they examined). The only exception is Nelson et al.'s (1987) finding that meaning set size affected stem completion (albeit with ending stems) and not word fragment completion.

In our three experiments word fragment and word stem completion were directly compared with other variables held constant. Under these conditions, we observed no differences in patterns of priming on word fragment and word stem completion tests across manipulations of material (words or pictures), levels of processing (orthographic or semantic), or delay (up to a week). In addition, Rajaram and Roediger (1992) conducted another experiment directly comparing stem and fragment completion (and other measures) over several variables and found no differences among these measures. Of course, we cannot extrapolate from these results that word fragment and word stem completion tests will not differ when other variables are manipulated, but the present results and those of Weldon et al. and Rajaram and Roediger (1992) leave us with the conclusion that the similarities between these tests are more important than their differences.

The apparent differences between word fragment and word stem completion tests that occurred in the earlier cross-experiment comparisons were likely due to confounded factors, especially the different types of words used on those tests. Because both tests are quite sensitive to format differences such as modality of presentation (picture or word and auditory or visual presentation of words), both tests seem to rely on perceptual procedures to produce priming (Roediger, 1990). In addition, both word fragment completion and word stem completion benefit from subjects' prior reading of words in relation to their generation from conceptual clues (see Blaxton, 1989, and Srinivas & Roediger, 1990, for relevant data from word fragment completion; Java & Gardiner, 1991 and McClelland & Pring, 1991, provide similar data for word stem completion; but, see Schwartz, 1989). Roediger, Weldon, and Challis (1989) argued that such comparisons were critical in establishing tests as perceptual or data-driven. Both word fragment completion and word stem completion seem to meet the criteria spelled out by Roediger, Weldon, and Challis (1989) for data-driven or perceptual tests.

Measuring Implicit Memory

Implicit memory tests are assumed to assess a form of memory that is different from that accessed by explicit tests of conscious recollection. A difficulty for this assumption is that subjects may well use explicit retrieval strategies to solve word fragments or stems and convert the tests into ones of cued recall. This is especially so in normal subjects; amnesic subjects and other memory-impaired populations are assumed to be unable (or at least much less able) to use explicit retrieval strategies on implicit tests; nonetheless, sometimes appeal is made to these subjects' "spared declarative memory" in explaining unwonted findings (e.g., Squire et al., 1987). Thus, more rigorous criteria are needed for establishing a test as implicit.

Schacter et al. (1989) suggested a logic they called the *retrieval intentionality criterion* whereby a variable (such as levels of processing) is manipulated and performance is compared on different versions of tests (explicit and implicit) using the same test cues. For example, Graf and Mandler (1984) showed sizeable levels-of-processing effects on cued recall with word stems as cues but no effect of this variable when the same word stems were presented with implicit test instructions. Because levels of processing had a large effect on explicit recall with word stems as cues but no effect on the implicit test with the same cues, Graf and Mandler concluded that the word stem completion test was not contaminated by explicit retrieval. We replicated their results in Experiment 1 and extended them to word fragment completion: Levels of processing of words had a large effect on explicit recall with word fragments as cues but no effect on priming in an implicit word fragment completion test. Therefore, word fragment completion also satisfies the retrieval intentionality criterion for an implicit test.

Ideally, one would like to find manipulations such that opposite patterns of results would occur on explicit and implicit versions of tests when test cues and all other conditions are held constant. For example, one might hope to find a picture superiority effect on explicit word fragment or stem cued-recall tests. However, given the nature of the word stem and word fragment tests, such an outcome is unlikely. Weldon et al. (1989) argued that explicit recall with word fragments or stems as cues involved both components of a perceptual test (to resolve the degraded word) and then a conceptual component (because explicit retrieval is assumed to be meaning-based; see Roediger, Weldon, & Challis, 1989).

Jacoby (1991) has argued that "task performance always represents a blend of automatic and intentional processes" (p. 514) and advocates a new process dissociation procedure to disentangle these processes, to more firmly attribute some effects to automatic responding and others to controlled responding. This procedure may prove generally useful, but the present results would seem to argue that (at best) only small aspects of the priming effects observed in word stem and word fragment completion tests arise from conscious recollection. If subjects were engaging in conscious recollection to any appreciable extent, then positive levels-of-processing effects would be observed for verbal materials (rather than their complete absence, as in the present data).

Levels-of-Processing Effects on Implicit Tests

Roediger, Weldon, and Challis (1989) argued that the standard levels-of-processing manipulation is one of a set of converging operations that can be used to distinguish perceptual from conceptual memory tests; this variable should affect conceptual tests, whether explicit or implicit, but have little or no effect on implicit perceptual tests. Level of processing does have an effect on conceptual implicit tests, such as generating words belonging to categories (Hamann, 1990; Srinivas & Roediger, 1990), as well as on explicit conceptual tests, such as recall or recognition. However, Challis and Brodbeck (1992) pointed out that there are usually small levels-of-processing effects on implicit perceptual tests such as word fragment or word stem completion. Although often not significant in a particular experiment, levels-of-processing effects have occurred numerically in virtually all prior experiments. The effect seems small but real. In their experiments, Challis and Brodbeck found large levels-of-processing effects when the manipulation was made with a between-subjects design; they found smaller but significant effects with a within-subjects design when items accorded a particular level of processing were blocked together. In their experiments, the only conditions in which no reliable effect of level of processing occurred on an implicit word fragment completion test were when items were randomly presented during study.

At one level, our results do not accord with Challis and Brodbeck's (1992) analysis. We used blocked presentation of the items in the various orienting task conditions in Experiment 1 yet still found no levels-of-processing effect on either implicit test following study of words. However, there may be a good reason for this: Although the manipulation of orienting task was blocked, presentation of words and pictures occurred randomly within each block of processing trials. This may have affected the encoding strategies subjects used. The conditions that sometimes produce small levels-of-processing effects on perceptual implicit memory tests remain to be determined. The present data show that strong dissociations can be obtained on implicit and explicit tests with the same test cues as a function of levels of processing. Furthermore, Brown and Mitchell (1992) concluded from a meta-analysis of this literature that the magnitude of levels-of-processing effects on implicit tests were generally uncorrelated with the effect on explicit tests. Thus, the levels-of-processing effect (or its absence) with words still seems a useful way to distinguish perceptual implicit tests from other, conceptual tests, whether implicit or explicit (Hamann, 1990).

The effect of manipulating orienting tasks for pictorial stimuli produced an interesting effect on both implicit tests: When subjects saw a picture in the graphemic condition, they were required to form an image of the word naming the picture and to count the ascending and descending letters in the word. Apparently this manipulation engaged central perceptual processes that are used in actually reading the word, so that greater priming occurred in the graphemic condition than in the pleasantness rating condition. The greater effect of word imaging than of pleasantness rating with pictures on a verbal implicit test fits well with theories arguing that priming in these tests is due to perceptual or lexical processes engaged during the time of study (e.g., Roediger, 1990; Tulving & Schacter, 1990; Weldon, 1991) but makes little sense from other perspectives. Semantic processing is typically thought to be deeper and hence better for retention on tests of memory, but in the present case, the task of imaging a word while studying a picture led to greater transfer on the verbal implicit tests than did prior study of pictures under semantic processing conditions. This outcome shows, once again, that types of processing are not inherently good or poor for all memory tests; rather, the efficacy of various orienting tasks depends critically on the nature of task requirements in the test and the appropriateness of prior processing for the test (Morris et al., 1977; Roediger, 1990).

The finding just described may mean that implicit perceptual tests could be a convenient vehicle for studying mental imagery. Jacoby and Witherspoon (1982) reported that spelling of auditorily presented words produced more priming on a perceptual identification test than did simply hearing the word, which may be attributable to imagining the word while spelling it (or at least to involving some form of graphemic processing). Donnelly (1988) produced a similar effect in primed word fragment completion. Similarly, Roediger and Blaxton (1987b, Experiment 1) gave subjects words auditorily and asked them to imagine what the word would look like either in a particular typeface or hand printed. Subjects later showed more priming on a word fragment completion test after this imagery manipulation than when they heard the words with no instructions. These results further show that perceptual implicit tests may be used to study mental imagery. Curiously, the published results to date have all used verbal tasks that, on the face of it, would not seem to be optimal for studying imagery. However, recently, McDermott (1991) has found about 10% greater priming from words when subjects imaged their referents than from words whose referents were not imaged on an implicit test involving naming of picture fragments.

Forgetting on Implicit Memory Tests

Although previous research has led to the conclusion that priming decays more rapidly on the word stem than the word fragment completion test, there was no evidence for this claim in our Experiments 2 and 3 when other factors were held constant. In the only other direct comparison experiment like ours of which we are aware, Gibson (1989, Experiment 4) reached a similar conclusion.[2]

Why are our results different from those reported in the literature? Without a good deal of further experimental work, there is no sure answer to this question. One possibility is that word frequency or familiarity plays a role, because this variable has been confounded with tests in most prior work. Researchers testing word fragment completion have used long, low-frequency words, whereas those using word stem completion have used short, high-frequency words. We did find greater priming from low- than from high-frequency words in Experiment 3, replicating MacLeod (1989). The effects were not large, but in relation to other studies, we probably used more homogeneous words because we also equated items in the different frequency conditions for their length and their imagery values. Others have not taken these extra steps and therefore may have confounded frequency with other basic variables such as length. In fairness, it is difficult to conduct experiments directly comparing frequency without confounding the critical comparison with some other variable. The word stem completion results in the bottom part of Table 4 are a case in point. Interpretation of priming scores is clouded by the fact that the base rates for low-frequency words are so much lower than for high-frequency words. When subjects receive the word stem, they are more likely to complete it with a high- than a low-frequency word,

---

[2] While this article was in press, we became aware of another experiment directly comparing loss of priming on word fragment and word stem completion tests. Craik, Moscovitch, and McDowd (1992, Experiment 2) showed equivalent priming on the two tests that did not decline over a 24-hr period.

thereby causing the inequality in base rates. Correcting this problem (e.g., by adding a letter to the stems for low-frequency words) would then only confound some other factor. Clearly, the difficulties in determining the amount of priming from various types of words and test cues are great, but most evidence indicates greater priming from low- than from high-frequency words.

Implications for a Generation/Recognition Model

Jacoby and Hollingshead (1990) have forwarded a generate/recognize model of performance for cued recall (see Humphreys, Bain, & Pike, 1989, for a similar approach). For example, in explicit recall with word fragments as cues, subjects must first try to generate a word to the fragment cue (presumably just as they would on an implicit test) and then try to recognize the word as having occurred in the list. With fragments or stems as cues, the generation process is greatly determined by perceptual factors in producing a word to the impoverished perceptual cue, whereas the recognition process is affected by conceptual factors (such as levels of processing). The assumptions can provide a plausible amount of our data in Tables 1 and 2 and the data of others (e.g., Nelson et al., 1987; Nelson, Keelean, & Negrao, 1989).

Briefly, words produce more priming than do pictures on the implicit tests because they aid the process of producing (generating) a word to the perceptual clue. Levels of processing of words at study affect performance on the explicit tests because of the sensitivity of the recognition component to conceptual factors.

For pictures the case is more complex but can still be understood in terms of the generate/recognize model. Pictures studied under semantic orienting conditions produced little priming on the verbal implicit tests, but when subjects studied pictures under conditions that encouraged them to image their names, priming increased. This enhanced generation of words at test from imagining their names during study also accounts for why performance was so good even on explicit tests for this study condition: The generation process in explicit retrieval was greatly enhanced in relation to the case in which subjects judged pleasantness of pictures.

Finally, Jacoby and Hollingshead (1990) predicted that when fragments or stems have only a single solution, performance on explicit versions of the test should not exceed that on implicit tests (generating and recognizing should not be greater than just generating). The prediction generally holds for the word fragment completion results in Tables 1 and 2. (The results from the word stem completion test are not relevant in our experiments, because the stems used were not unique.)

In general, results from Experiments 1 and 2 accord with Jacoby and Hollingshead's (1990) generation/recognition model, with the added assumptions that generation from single word fragments is largely guided by perceptual (data-driven) factors and that recognition is mostly determined by conceptual factors.

Conclusion

On the basis of the present results, the implicit memory tests of word stem and word fragment completion seem quite similar. Both tests show greater priming from words than from pictures, no levels-of-processing effects for words, and roughly similar rates of forgetting. In addition, both tests meet the retrieval intentionality test suggested by Schacter et al. (1989) for classification as implicit tests. No doubt future research will uncover differences between the tests (perhaps from manipulations that would draw attention either to the first letters of a word or to its overall appearance), but from the present data the similarities between the tests seem quite impressive.

References

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 657–668.

Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 404–416.

Brown, A. S., & Mitchell, D. B. (1992). *Levels of processing in implicit memory: A re-evaluation.* Unpublished manuscript.

Butters, N., Heindel, W. C., & Salmon, D. P. (1990). Dissociation of implicit memory in dementia: Neurological implications. *Bulletin of the Psychonomic Society, 28,* 359–366.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116–131.

Challis, B. H., & Brodbeck, D. R. (1992). Level of processing affects priming in word fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 595–607.

Craik, F. I. M., Moscovitch, M., & McDowd, J. (1992). *Contributions of surface and conceptual information to performance on explicit and implicit memory tasks.* Manuscript submitted for publication.

Donnelly, R. E. (1988). *Priming across modality in implicit memory: Facilitation from auditory presentation to visual test of word-fragment completion.* Unpublished doctoral dissertation, University of Toronto, Toronto, Ontario, Canada.

Durgunoglu, A. Y., & Roediger, H. L., III. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language, 26,* 377–391.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto word pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation, 14,* 375–399.

Gibson, J. M. (1989). *An investigation into the decline in priming of word completion: A test of the number of completions hypothesis.* Unpublished doctoral dissertation, Rice University, Houston, TX.

Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior, 23,* 553–568.

Graf, P., Mandler, G., & Haden, P. (1982). Simulating amnesic symptoms in normal subjects. *Science, 218,* 1243–1244.

Graf, P., Shimamura, A. P., & Squire, L. R. (1985). Priming across modalities and priming across category levels: Extending the domain of preserved function in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 386–396.

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 164–178.

Greene, R. L. (1986). Word stems as cues in recall and completion tasks. *The Quarterly Journal of Experimental Psychology, 38A,* 663–673.

Hamann, S. B. (1990). Level-of-processing effects in conceptually driven implicit tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 970–977.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

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

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic and intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests of memory. *Journal of Memory and Language, 29,* 433–454.

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology, 32,* 300–324.

Java, R. I., & Gardiner, J. M. (1991). Priming and aging: Further evidence of preserved memory function. *American Journal of Psychology, 104,* 89–100.

Komatsu, S. I., & Ohta, N. (1984). Priming effects in word fragment completion for short- and long-retention intervals. *Japanese Psychological Research, 26,* 194–200.

Kucera, M., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

MacLeod, C. M. (1989). Word context during initial exposure influences degree of priming in word fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 398–406.

McClelland, A. G. R., & Pring, L. (1991). An investigation of cross-modality effects in implicit and explicit memory. *The Quarterly Journal of Experimental Psychology, 43A,* 19–33.

McDermott, K. B. (1991). [Effects of imagery instructions on primed picture fragment naming]. Unpublished raw data.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519–533.

Neill, W. T., Beck, J. L., Bottalico, K. S., & Molloy, R. D. (1990). Effects of intentional versus incidental learning on explicit and implicit tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 457–463.

Nelson, D. L., Canas, J. J., Bajo, M. T., & Keelean, P. D. (1987). Comparing word fragment completion and cued recall with letter cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 542–552.

Nelson, D. L., Keelean, P. D., & Negrao, M. (1989). Word-fragment cuing: The lexical search hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 388–397.

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

Rajaram, S., & Roediger, H. L. (1992). *Direct comparison of four implicit memory tests.* Unpublished manuscript.

Random House college dictionary (Rev. ed.). (1984). New York: Random House.

Rappold, V. A., & Hashtroudi, S. (1991). Does organization improve priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 103–114.

Roediger, H. L., III. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

Roediger, H. L., III, & Blaxton, T. A. (1987a). Effects of varying modality, surface features, and retention interval on priming in word fragment completion. *Memory & Cognition, 15,* 379–388.

Roediger, H. L., III, & Blaxton, T. A. (1987b). Retrieval modes produce dissociations in memory for surface information. In D. Gorfein & R. R. Hoffman (Eds.), *Memory and cognitive processes: The Ebbinghaus centennial conference* (pp. 349–379). Hillsdale, NJ: Erlbaum.

Roediger, H. L., III, Srinivas, K., & Weldon, M. S. (1989). Dissociations between implicit measures of retention. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 67–84). Hillsdale, NJ: Erlbaum.

Roediger, H. L., III, Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–14). Hillsdale, NJ: Erlbaum.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1–17.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schacter, D. L. (1990). Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. *Annals of the New York Academy of Sciences, 608,* 543–567.

Schacter, D. L., Bowers, J., & Booker, J. (1989). Intention, awareness and implicit memory: The retrieval intentionality criterion. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 47–65). Hillsdale, NJ: Erlbaum.

Schoen, L. M., Ciofalo, E., & Rudow, E. (1989). Anagram versus word fragment solution: A comparison of implicit memory measures. *Bulletin of the Psychonomic Society, 27,* 551–552.

Schwartz, B. L. (1989). Effects of generation on indirect measures of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1119–1128.

Shimamura, A. P. (1986). Priming effects in amnesia: Evidence for a dissociable memory function. *Quarterly Journal of Experimental Psychology, 24,* 490–504.

Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 223–239.

Smith, E. R., & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology, 14,* 223–239.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General, 109,* 373–392.

Snodgrass, J. G., & Vanderwart, M. A. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

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.

Srinivas, K., & Roediger, H. L., III. (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory and Language, 29,* 389–412.

Tenpenney, P. L., & Shoben, E. J. (1992). Component processes and the utility of the conceptually-driven/data-driven distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 25–42.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Tulving, E., Hayman, C. A. G., & Macdonald, C. A. (1991). Long-lasting perceptual priming and semantic learning in amnesia: A case experiment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 595–617.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 336–342.

Warrington, E. K., & Weiskrantz, L. (1978). Further analysis of the prior learning effect in amnesic patients. *Neuropsychologia, 16,* 169–177.

Webster's new international dictionary, unabridged (2nd ed.). (1934). Springfield, MA: Merriam-Webster.

Weldon, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 526–541.

Weldon, M. S., & Roediger, H. L., III. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition, 15,* 269–280.

Weldon, M. S., Roediger, H. L., III, & Challis, B. H. (1989). The properties of retrieval cues constrain the picture superiority effect. *Memory & Cognition, 17,* 95–105.

Appendix A

Items Used in Experiment 1

Pictures were those from Snodgrass and Vanderwart (1980) or Weldon and Roediger (1987). Word stems were the first three letters of each word.

Item	Fragment	Item	Fragment	Item	Fragment
ambulance	a__ula_c_	giraffe	__r_f_e	scissors	__isso__
arrow	_r__w	grasshopper	_ra_s_op___	screwdriver	s__ew_ri_e_
ashtray	__ht_ay	headphones	he__p_on_s	snowman	_no__an
backgammon	__ckga_mo_	intestines	in__stin_s	squirrel	s_ui_re_
balloon	__l_oo_	lobster	l_bs___	strawberry	s_ra___r_y
barrel	ba___l	microscope	_i_ros_op_	submarine	__bm_ri_e
bicycle	__c__le	motorcycle	_oto__y___	suitcase	__itc_se
binoculars	_ino__la_s	mountain	m_un__in	suspenders	_us_end_r_
butterfly	_utt___l_	mushroom	m__h_o_m	sweater	_w_ate_
camera	_ame_a	octopus	__topu_	tambourine	t__b__rin_
canoe	___oe	paintbrush	p___tb_us_	telescope	__le_c_pe
carrot	c___ot	parachute	_ar_ch_t_	thermometer	_her__m_te_
caterpillar	__ter_il__r	peacock	_e_co_k	thimble	_hi_b_e
corkscrew	_ork_cre_	pencil	_e_c_l	tornado	_orn_d_
diamond	_ia__nd	pineapple	pi___p__e	typewriter	t__e_rit__
elevator	__ev_to_	pretzel	_ret__l	unicorn	_n_cor_
envelope	_nve___e	raccoon	r___oo_	violin	_io__n
escalator	_sca__to_	refrigerator	_efri___ato_	watermelon	__te_m__on
feather	f_ath__	rhinoceros	__in_ce_os	wheelchair	w_e_lc_a__
fireplace	_ire_la_e	sandwich	s__d_i_h	windmill	_ind_il_

Appendix B

Detailed Statistical Analyses of Experiment 1

Analyses were based on differences between studied and nonstudied items. In the explicit retrieval conditions, these scores are the proportion of items recalled minus the proportion of false alarms to nonstudied items. In the implicit tests, scores are priming (studies − nonstudied completions).

The experiment was a Retrieval (implicit vs. explicit) × Test (fragments vs. stems) × Study (intentional vs. incidental) × Levels of Processing (graphemic vs. semantic) × Format (pictures vs. words) mixed-design ANOVA with repeated measures on levels of processing and format. A 5-way ANOVA revealed no significant interactions of the intentional/incidental factor with any of the others. The rest of the analyses were collapsed over the study variable. Furthermore, two 3-way ANOVAs were completed, one for the implicit test conditions and one for the explicit test conditions. Summaries of these analyses are presented below. All reported effects were significant at the .05 level. Nonsignificant effects are not reported. *T* denotes the retrieval test variable (implicit vs. explicit), *L* the levels-of-processing variable (graphemic vs. semantic), *C* the test cue variable (fragments vs. stems), and *F* the study format variable (pictures vs. words).

Effect	F	MS_e	Comment
			Four-way ANOVA
T	183.79[a]	.070	The level of performance in implicit conditions (.12) was lower than in explicit conditions (.36).
L	16.76[a]	.028	Semantic processing (.26) produced better performance than graphemic processing (.22).
F	18.23[a]	.032	Words (.27) produced better performance than pictures (.22).
TC	12.35[a]	.070	The difference between implicit and explicit retrieval conditions was bigger for stems (.10 and .39, respectively) than for fragments (.15 and .32, respectively).
TL	61.76[a]	.028	In the implicit retrieval condition, graphemic processing produced better performance than semantic processing (.14 and .10, respectively), but in the explicit retrieval condition, graphemic processing produced worse performance than semantic processing (.29 and .42, respectively).

TF	50.46[a]	.032	In the implicit retrieval condition, pictures produced worse performance than words (.06 and .19, respectively), but in the explicit retrieval condition, pictures produced slightly better performance than words (.37 and .34, respectively).
CL	8.36[a]	.028	There was a smaller difference in performance between the graphemic and semantic levels of processing for fragments (.23 and .24, respectively) than for stems (.21 and .28, respectively).
CF	11.54[a]	.032	There was a much larger difference in performance between pictures and words for fragments (.28 and .19, respectively) than for stems (.25 and .24, respectively).
LF	120.27[a]	.210	Pictures produced better performance than words (.25 and .19, respectively) for the graphemic processing condition, but worse performance (.19 and .34, respectively) for the semantic processing condition.
TLF	58.22[a]	.210	Under implicit retrieval conditions, words produced better performance than pictures under both graphemic (.19 and .10, respectively) and semantic (.19 and .02, respectively) levels of processing. Under explicit retrieval conditions, pictures produced better performance than words (.40 and .19, respectively) under graphemic processing instructions, but pictures produced worse performance than words (.35 and .49, respectively) under semantic processing instructions.
CLF	4.47[a]	.021	On word fragments, words produced the same performance as pictures (.23 and .23, respectively) after graphemic processing, but words produced much better performance than pictures (.33 and .16, respectively) after semantic processing. On word stems, words produced worse performance than pictures (.15 and .26) after graphemic processing, but words produced better performance than pictures (.35 and .22) after semantic processing.
TCLF	8.23[a]	.021	See Tables 1 and 2 for means. Basically, words always produced better performance than pictures in the implicit conditions, whereas in the explicit conditions, words produced better performance than pictures with semantic processing instructions, and pictures produced better performance than words with graphemic processing instructions. The sizes of these effects were slightly different for fragments and stems, with the difference between pictures and words being larger for fragments than for stems (with the exception of graphemic processing instructions under explicit retrieval instructions).
			Three-Way ANOVA: Implicit tests
C	5.72[b]	.052	Fragments (.15) produced more priming than stems (.10).
L	8.15[b]	.024	Graphemic processing (.15) produced more priming than semantic processing (.10).
F	62.50[b]	.033	Words (.19) produced more priming than pictures (.06).
LF	5.87[b]	.020	Graphemic processing produced more priming than semantic processing with pictures (.10 and .02, respectively) but not with words (.19 in both cases).
			Three-way ANOVA: Explicit Tests
C	6.74[b]	.089	Stems (.39) produced better recall than fragments (.32).
L	63.20[b]	.031	Semantic processing (.42) produced better recall than graphemic processing (.29).
F	4.16[b]	.030	Pictures (.37) produced slightly greater recall than words (.34).
CL	5.40[b]	.031	The difference between graphemic and semantic processing was smaller for fragments (.28 and .37, respectively) than it was for stems (.31 and .48, respectively).
CF	11.91[b]	.031	Words produced slightly better performance than pictures with fragments (.33 and .31, respectively), but the opposite pattern emerged with stems (.35 and .44, respectively).
LF	164.35[b]	.022	Semantic processing produced better performance than graphemic processing with words (.49 and .19, respectively), but the pattern was reversed with pictures (.35 and .40, respectively).
CLF	11.80[b]	.022	For fragments, semantic processing produced better performance than graphemic processing with words (.44 and .22, respectively), but the pattern was reversed for pictures (.29 and .33, respectively). For stems, the same pattern emerged, but the differences were bigger. Semantic processing produced better performance than graphemic processing with words (.54 and .16, respectively), but the pattern was reversed for pictures (.41 and .46, respectively).

[a] *df* = 1,236. [b] *df* = 1,118.

(Appendix C follows on next page)

Appendix C

Words, Anagrams, and Fragments Used in Experiment 2

Word	Anagram	Fragment
almanac	cmalnaa	al__n_c
anatomy	antmoya	_na__my
anchor	cnhroa	_n__or
antenna	ntanena	_n__nna
anybody	bdyoayn	an___dy
apple	elpap	_p_le
arrow	rwaor	_r__w
ashtray	yrashat	__ht_ay
asylum	symlua	a__l_m
avocado	acdoavo	_v_cad_
bakery	krbaey	b_k__y
balloon	bnlaloo	__l_oo_
barrel	rbrale	ba__e_
bashful	sahulbf	b_sh_u_
begonia	oianegb	_eg_ni_
beetle	etlebe	b_e_le
bicycle	ycielbc	__cy_le
bourbon	rounbbo	_ou__on
boyhood	yhobodo	_oyh__d
bravado	odbrvaa	br__a_o
bullock	kuclobl	bu_l__k
bureau	rbuaue	bu_e_u
button	utobtn	_ut__n
cabaret	btcarea	c__ar_t
carrot	racotr	c___ot
cathode	thacdoe	c_th__e
cavalry	vlayrca	c_va__y
celery	lercye	c_le_y
chassis	hascssi	__ass_s
chicken	kcnehci	c_ic_e_
cloud	lduco	c__ud
cholera	olercha	c_o__ra
cinder	nderic	ci_de_
climate	mtlcaie	_l_m_te
cobbler	berlboc	c__ble_
cocaine	ocneiac	c_c_in_
comma	macmo	c__ma
conifer	ronifce	c_nif__
copper	pocerp	c_pp__
cotton	onctot	_ot__n
country	tuncory	__u_t_y
crevice	vcrecei	_rev_c_
curator	rtrouca	cu___or
death	htdae	d_a_h
dolphin	phodlin	do__h_n
factory	caoryft	f__t_ry
fascism	safsmic	f__c_sm
finger	nifger	f__g_r
fishing	ginshif	f_sh_n_
flower	lwofre	f__we_
freckle	crefkle	f_ec__e
fusion	suonif	fu_i_n
giraffe	fgreaif	__r_f_e
glove	elgov	gl_v_
hanger	erhgna	ha_g_r
heart	threa	h_a_t
hemlock	mlcokeh	he__o_k
house	souhe	h_us_
inertia	iatreni	_ne_t_a
lanolin	nalolni	__no_in
ladybug	gdlabyu	la_y__g
lariat	rliata	l_r_at
lettuce	tteulce	_et_u_e
liberty	bteryli	l_b_r_y
lineage	nilegea	l_ne__e
lithium	thmulii	__th_um
maggot	gamotg	m_gg__
marble	rmaleb	ma__l_
mascara	casmraa	m_s_a_a
meadow	dwomae	_e_d_w
menthol	lmntheo	m__t_ol
midget	dmgite	mi_g_t
mystery	myrstye	_ys__ry
nickel	knicle	_ic_e_
obelisk	kisbloe	ob__i_k
octopus	cotuspo	o_t__us
office	fecofi	of__c_
orange	rnageo	o__ng_
peacock	eapkcoc	pe_c_c_
penguin	uipnegn	pe__ui_
pepper	rppepe	_ep__r
petunia	etniaup	pe_u_i_
pumpkin	kimuppn	pu_p_i_
pigment	gipenmt	p_gm__t
pitcher	ercthip	_it_h_r
plague	guapel	p___u_
pretzel	ztperle	_ret__l
proverb	bporvre	p_o___b
purple	rpupel	_urp__
raccoon	ocoranc	r___oo_
radish	drashi	_a_ish
rainbow	ainwobr	r___b_w
rabbit	bbtiar	__b_it
recluse	eclsrue	r_cl_s_
refrain	nifarer	r_f_ai_
rooster	sdotorer	r_o__e_
rotunda	ortunda	ro___da
rubella	ulalbre	ru__l_a
saliva	alsiva	_al_v_
sandal	dnalas	s_n_al
school	cooslh	_c_o_l
season	ssonae	s_a_on
shadow	whsado	_h_do_
sheriff	erifhsf	s____ff
shovel	lsohve	_h_v_l
skirt	krits	sk_r_
slipper	pprelis	s_ip_e_
smile	meils	_mil_
snake	sknae	_n_ke
snowman	wonsnma	_no_ma_
soldier	erlodsi	s_ld__r
spatula	patslua	_p_tu_a
spoon	nopos	_p_on
spider	presid	_p_d_r
stone	nsteo	_t_ne
summer	msumre	su_m__
surgery	urgsrye	s__ge_y
swahili	sawilih	s_ah__i
sweater	retawse	_w_ate_
teacher	aechtre	_e_ch_r
theorem	hoetemr	_h_o_em
thimble	himtelb	_hi_b_e
tornado	rotadon	_orn_d_
tragedy	ragtyde	t_ag___
treason	rsteaon	_re__on
trumpet	tepmutr	___mpe_
turkey	urtyek	t_r__y
unicorn	niuronc	_n_cor_
vacuum	uuavcm	_acu__
valium	lvaumi	v_li__
vanilla	ilanavl	v_n_l__
viceroy	viecoyr	_ic_ro_
violin	noiliv	_io__n
vulture	ulvurte	v_lt__e
whiskey	kyhiswe	w_i__ey

Appendix D

Words and Fragments Used in Experiment 3

Word stems were the first three letters of the words. H and L represent high- or low-frequency words, respectively.

Item	Frequency	Fragment	Item	Frequency	Fragment
absence	H	a_s_n_e	heredity	L	_e_ed_t_
ability	H	_bi__ty	history	H	h__to__
account	H	ac___nt	husband	H	_u_b__d
admiral	L	_dm__al	impetus	L	_m_e_us
advantage	H	ad_an__g_	incline	L	_n_li_e
adultery	L	_d_lt_r_	inspire	L	i_sp_r_
alias	L	_l_as	industry	H	_nd___ry
alligator	L	a___ga__r	knowledge	H	_n_w___ge
amount	H	a__un_	language	H	l_n_u__e
ancient	H	_nc_en_	machine	H	_ac__ne
anecdote	L	_n_c_ot_	maiden	L	m_i__n
animal	H	a_i_a_	malaria	L	m_l_ri_
appearance	H	a_p_ar__c_	material	H	m_ter___
armadillo	L	_rm_di___	measure	H	__as_r_
ashore	L	a_ho_e	memory	H	me__r_
attitude	H	at_i_u__	mermaid	L	m__m_i_
automobile	H	_ut_m___l_	method	H	_e_h_d
bacteria	L	b_c___i_	million	H	m__l_on
balance	H	_ala__e	mischief	L	_is_hi__
barnacle	L	b_r_ac_e	modern	H	m_d__n
beauty	H	b_a_t_	molecule	L	m_le_u__
bedroom	H	_e_ro_m	monarch	L	m__ar_h
beggar	L	b___ar	murder	H	m_rd__
behold	L	_eh__d	nutrient	L	n_t_i_n_
belief	H	_e_i_f	oblivion	L	_bl_v__n
betray	L	_et__y	occasion	H	_c_as__n
blessing	L	ble_si_g	officer	H	__fic__
business	H	_u_in__s	opportunity	H	_p_or_u_it_
butcher	L	_u_ch_r	painting	H	_a_n_i_g
cannon	L	c__n_n	panorama	L	_an_r_m_
captain	H	_ap__in	patient	H	_at_e_t
caravan	L	c_r__an	perfume	L	_e_fu_e
chimney	L	_hi_ne_	picture	H	p_ct___
college	H	c_l_eg_	platform	H	p_at___m
crucifix	L	c__c_fi_	pleasure	H	p_ea___e
culture	H	c__t_re	position	H	p_s__io_
damsel	L	d_ms__	prisoner	L	p_i_on__
deceive	L	_e_e_v_	product	H	_ro_u_t
development	H	__ve_o_me__	projectile	L	_ro_e_t_l_
design	H	d_si__	publish	L	_u_lis_
diamond	L	__a__nd	railroad	H	rai_r__d
digest	L	__g_st	regency	L	r__enc_
disease	H	_is_a_e	sailor	L	s__lo_
doctor	H	d__t_r	scarlet	L	_ca__et
dragon	L	d__go_	sedative	L	s_d_ti__
engine	H	__gin_	senate	H	s_na__
epitaph	L	e_it__h	shelter	H	s_el_e_
evergreen	L	_ve_g__e_	silence	H	s_le_c_
evidence	H	e_i_en__	slumber	L	s__mb_r
expression	H	e__re_si_n	southern	H	s__th__n
failure	H	f_i_ur_	speech	H	s_ee__
famous	H	_a_ou_	spinach	L	s_in__h
feather	L	fe_t__r	sprinkle	L	s_r_nk__
figure	H	f__ur_	stanza	L	_ta__a
flourish	L	_l_ur__h	surround	L	s_rr__n_
forest	H	_or_s_	thicket	L	t_ic__t
franchise	L	_r_n_h_s_	torment	L	t_rm_n_
friend	H	f___nd	traitor	L	t_ai_o_
function	H	_u_ct_on	trouble	H	t_o_b_e
gauntlet	L	_au_t_e_	typhoon	L	__pho_n
glitter	L	g_it_e_	twinkle	L	t_i_k_e
goddess	L	g_d_es_	university	H	u__ve__i_y
health	H	h_a_t_	vanish	L	v__is_
helmet	L	h_lm__	volume	H	__lu_e

Received December 27, 1991
Revision received March 17, 1992
Accepted March 18, 1992 ■