Journal of Experimental Psychology:
Learning, Memory, and Cognition
1996, Vol. 22, No. 1, 132–142

0278-7393/96/$3.00

Word Frequency Effects on Recall, Recognition, and Word Fragment Completion Tests

Colin M. MacLeod and Kristina E. Kampe
University of Toronto, Scarborough Campus

In 3 experiments, the effect of word frequency on an indirect word fragment completion test and on direct free-recall and Yes–no recognition tests was investigated. In Experiment 1, priming in word fragment completion was substantially greater for low-frequency words than for high-frequency words, but free recall was unaffected. Experiment 2 replicated the word fragment completion result and showed a corresponding effect in recognition. Experiment 3 replicated the low-frequency priming advantage in word fragment completion with the set of words that P. L. Tenpenny and E. J. Shoben (1992) had used in reporting the opposite pattern in word fragment completion. Using G. Mandler's (1980) dual-process theory, the authors argue that recognition and word fragment completion tests both rely on within-item integration that influences familiarity, whereas recall hinges on elaboration that influences retrievability.

One of the earliest empirical observations in cognitive psychology was made by Cattell (1886b). He demonstrated that the frequency of occurrence of a word in a language affects even the most basic processing of that word, its speed of recognition. People do not process *aardvark* or *kumquat* as quickly—or indeed in the same way—as they process *horse* or *apple.* That same year, in attempting to explain why words are read aloud faster than the corresponding objects or their properties can be named aloud, Cattell (1886a) suggested that words were processed automatically, whereas other stimuli were not. With a slight modification, this same explanation could be extended to the word frequency effect: The degree of automaticity of processing words decreases as their frequency in the language decreases. Automaticity is a direct function of experience.

Since Cattell's pioneering work, word frequency has been a persisting subject of study for investigators concerned with the identification or recognition of words (e.g., Becker, 1979; Monsell, Doyle, & Haggard, 1989; Morton, 1969; Solomon & Postman, 1952). Numerous potential confounds have been eliminated, and word frequency has remained as a potent variable. In general, these studies have shown that high-frequency words are identified faster than low-frequency words (e.g., Howes & Solomon, 1951), although processes beyond identification may also be affected (Monsell et al., 1989). All of these results strongly suggest that a word's frequency is part of its representation in memory.

Because memory must underlie this identification difference, investigators studying memory have continued to be interested in word frequency effects (e.g., Hall, 1954; Jacoby & Dallas, 1981; Peters, 1936; Scarborough, Cortese, & Scarborough, 1977; Sumby, 1963). In most studies of word frequency effects in memory, researchers have used conventional recognition or recall tests as the dependent measure, tests now classified as direct tests because they require awareness of retrieval from memory (see Richardson-Klavehn & Bjork, 1988; Schacter, 1987). Only very recently has exploration begun into word frequency effects on indirect tests, which can be performed without awareness of retrieval from memory.

Word Frequency Effects on Direct Tests of Memory

The commonly accepted wisdom for direct tests can be seen in two summary statements made a decade apart. Kintsch (1970, p. 277) wrote that "while it is known that high-frequency words tend to be recalled better than low-frequency words (Hall, 1954) the opposite relationship holds for recognition (Schwartz & Rouse, 1961; Gorman, 1961; Shepard, 1967)." Mandler (1980, p. 267) made very much the same claim:

> Subjects are more likely to recognize a low frequency word as having been previously presented than a high frequency word. Free recall follows the more expected pattern, of course, in that high frequency items are easier to recall than low frequency ones.

This pattern can be taken as evidence of a dissociation between recall and recognition, a finding of considerable interest to theorists (e.g., Gillund & Shiffrin, 1984; Mandler, 1980).

In fact, the dissociation is a little more complicated: There are important constraints on both the recall and recognition results. Gregg (1976) highlighted the recall constraint in his review of the literature, and it is a focal point in the Gillund and Shiffrin (1984) article. High-frequency words are better

Colin M. MacLeod and Kristina E. Kampe, Division of Life Sciences, University of Toronto, Scarborough Campus, Scarborough, Ontario, Canada.

This research was supported by Grant A7459 from the Natural Sciences and Engineering Research Council of Canada. Experiments 1 and 2 formed Kristina E. Kampe's senior honor's thesis. Experiments 1 and 2 were also reported at the annual meeting of the Psychonomic Society in Atlanta, Georgia, in November 1989. We thank Shelley Hodder and Haley Naor for collecting the data for Experiment 3.

Correspondence concerning this article should be addressed to Colin M. MacLeod, Division of Life Sciences, University of Toronto, Scarborough Campus, Scarborough, Ontario, Canada M1C 1A4. Electronic mail may be sent via Internet to macleod@lake.scar.utoronto.ca.

recalled than low-frequency words only for pure lists of either high or low word frequency, typically in a between-subjects design (e.g., Bousfield & Cohen, 1955; Shepard, 1967; Underwood, Ekstrand, & Keppel, 1965). In a within-subject design, in which a mixed list of both high- and low-frequency words is used, the word frequency effect in recall generally disappears, and low-frequency words are recalled as well as, or sometimes even a little better than, high-frequency words (e.g., Duncan, 1974; Gregg, 1976).

In two studies, this impact of design on the frequency effect in recall has been examined directly. Both Gregg, Montgomery, and Castano (1980) and Gillund and Shiffrin (1984) found the recall advantage for high-frequency words in pure lists but not in mixed lists. The same was true when trials to acquisition of the list, instead of recall after a single study trial, were examined (May & Tryk, 1970). Thus, what effect a word frequency manipulation should have on a direct test of recall hinges on the design of the experiment. Recall rarely shows the low-frequency advantage characteristic of recognition, except in a couple of cases in mixed lists (Duncan, 1974; May & Tryk, 1970).

With regard to recognition, the results are more consistent, at least for item recognition. In addition to the studies cited by Kintsch (1970), there are more recent studies that make the same point: Performance on direct item recognition tests is better for low-frequency words than for high-frequency words—whether in yes–no recognition (e.g., Balota & Neely, 1980) or in forced-choice recognition (e.g., Glanzer & Bowles, 1976)—even when the dependent measure is latency rather than the standard accuracy measure (Duchek & Neely, 1989). Indeed, Wallace, Sawyer, and Robertson (1978) even obtained this pattern without distractors; they instructed participants that all of the test items were in fact old but asked them to indicate only the ones they definitely remembered. Presumably, this low-frequency advantage in item recognition is related to the considerably greater increment in familiarity furnished by studying a low-frequency word versus a high-frequency word, as has been argued cogently by Mandler (1980, pp. 266–268).

Quite recently, Clark (1992; Clark & Burchett, 1994) has placed a qualification on the recognition result as well. His work replicated the low-frequency advantage in standard item recognition tests, but he reported that it reversed for associative recognition of pairs (Clark, 1992) or triples (Clark & Burchett, 1994). Associative recognition—that is, recognizing whether a test pair is an intact studied pair or the recombination of items from two studied pairs—is better for high-frequency words. Clark and Burchett also replicated the pure–mixed list effect in recall but showed that there was no such effect on either type of recognition.

In summary, then, direct tests show a complicated but reasonably consistent pattern. In free recall, there is a high-frequency advantage in pure lists but not in mixed lists, in which there usually is no frequency effect. The pure–mixed manipulation does not affect recognition. For item recognition, there is a low-frequency advantage; for associative recognition, there appears to be a weak high-frequency advantage.

Word Frequency Effects on Indirect Tests of Memory

There is less known about the effects of word frequency manipulations on indirect tests, partly because of the recency of their development. Two of the most widely recognized indirect tests (cf. Richardson-Klavehn & Bjork, 1988)—lexical decision and masked word identification—both consistently show better overall performance for high-frequency words than for low-frequency words (e.g., Duchek & Neely, 1989; Forster & Davis, 1984; Jacoby & Dallas, 1981; Kirsner, Milech, & Standen, 1983; Scarborough et al., 1977). However, overall performance is not the primary concern in studies in which indirect tests are used; rather, the major interest is in *priming*, the increment in performance accruing to a target item because of prior presentation of that item. Typically, the priming result is the reverse of the overall finding: Priming is greater for low- than for high-frequency words, a conclusion that also derives from the five studies just cited. This fact led Jacoby and Dallas to argue that a direct test (recognition) and an indirect test (masked word identification) actually showed the same pattern in response to word frequency "if one considers *change* in performance produced by prior study rather than absolute performance" (1981, p. 320). That is, both item recognition and priming in masked word identification showed a low-frequency advantage.

In a recent article, MacLeod (1989b) also observed more priming for low-frequency than for high-frequency words, this time on a word fragment completion test, another widely used indirect test. However, that study was designed to explore context effects on indirect tests and did not systematically manipulate word frequency. The observed frequency effect was based on an after-the-fact item analysis. Consequently, variables other than word frequency could have determined the outcome. Also, there was no direct test with which to contrast the indirect test results. One purpose of the present study was to examine these issues in detail.

The timeliness of the present study was also enhanced by the recent appearance of an article by Tenpenny and Shoben (1992). In their Experiment 3, they examined the impact of a word frequency manipulation on word fragment completion. What they found is rather surprising in light of the studies just described: They observed reliably more priming for high-frequency words than for low-frequency words. This result is at odds with the rest of the literature on word frequency effects on indirect remembering, including the two studies in which the same indirect test of word fragment completion was used (MacLeod, 1989b; Roediger, Weldon, Stadler, & Riegler, 1992), and therefore warrants further exploration. This provided a major motivation for our research: to resolve the nature of the word frequency effect on the word fragment completion test.

Word Frequency and Type of Test

The original aim of the present experiments was to investigate the dissociative effects of word frequency on different direct and indirect tests of memory. Word frequency is one of only a small set of variables known to dissociate recall and item recognition, two direct tests of memory. Word frequency also affects repetition priming in lexical decision and masked word identification, but in the same way as each other—more priming for low-frequency than for high-frequency words, analogous to the standard finding on a direct recognition test. Ultimately, though, a second aim emerged for the present study: to clarify the nature of the word frequency effect on the word fragment completion test in particular. After MacLeod (1989b) obtained results that suggested a low-frequency priming advantage in word fragment completion, Tenpenny and Shoben (1992, Experiment 3) had reported a high-frequency advantage, whereas Roediger et al. (1992, Experiment 3) had reported a low-frequency advantage. We investigated this discrepancy in Experiment 3.

In the three experiments reported in this article, word frequency was manipulated in mixed-list designs. In each of the first two experiments, we used a direct test of memory (free recall in Experiments 1 and 3 and item recognition in Experiment 2), and in all three experiments we used the same indirect test of memory (word fragment completion). This scheme had the advantages of permitting (a) replication of the novel word fragment completion results over the three experiments and (b) separate contrasts of the indirect word fragment completion test with the two most widely used direct tests: recall and recognition.

Experiment 1

In this first experiment, the expectation was that word frequency should affect free recall and word fragment completion differently. When using a mixed-list design, there should be no effect of word frequency on free recall, the direct test (e.g., Gillund & Shiffrin, 1984; Gregg et al., 1980). This, of course, simply would provide replication of an already well-replicated result, but with the same materials to be used in word fragment completion. Prediction of the effect of the word frequency manipulation on word fragment completion was not as straightforward. On the basis of the results from other indirect measures (e.g., Jacoby & Dallas, 1981) and two studies of word fragment completion (MacLeod, 1989b; Roediger et al., 1992, Experiment 3), low-frequency words should show more priming than high-frequency words in word fragment completion, the indirect test. However, the MacLeod analysis was post hoc, and the Roediger et al. result displayed only 6% more priming for low-frequency words (.15) than for high-frequency words (.09). Furthermore, the findings of Tenpenny and Shoben (1992, Experiment 3) lead to the opposite conclusion for word fragment completion: They observed 7% more priming for high-frequency words (.26) than for low-frequency words (.19). Regardless of the direction of the difference in priming in word fragment completion, however, the predicted result was a dissociation between the direct and indirect tests.

Method

Participants. The participants were 24 students from introductory psychology classes at the Scarborough Campus of the University of Toronto who received bonus points toward their grade for taking part in the experiment.

Apparatus. Stimulus presentation was controlled by a QuickBASIC program running on an IBM-compatible 286 computer with an amber monochrome monitor. Millisecond accuracy timing routines were taken from Graves and Bradley (1987, 1988). All materials were printed in regular lowercase font with 80 characters per line.

Materials. A pool of 96 critical words was selected from the Dahl (1979) norms. Half of the words were high in frequency (20 or more per million), and half were low in frequency (1 or fewer per million). In addition, a further set of 48 words was selected for use as filler material on the word fragment completion test. These filler items, which were not controlled for word frequency, are not included in the analyses in either experiment. The entire set of 144 words is presented in the Appendix. All were either nouns or adjectives and ranged in length between 6 and 10 letters.

For each of the words, a corresponding fragment was created; in the Appendix we indicate how a fragment was formed by underlining those letters omitted from each word. The particular fragments used in the first two experiments were chosen from a larger set examined in a pilot study. The criteria for choice were that (a) there be only one possible completion for a given fragment and (b) the likelihood of completing every fragment be about .20.

We intentionally equated baseline completion probability (without priming) for the high- and low-frequency words. Of the two possible strategies—similar deletion rules for letters in both types of items with a consequent difference in difficulty of completion, or different deletion rules leading to equivalent difficulty of completion—we favored the latter, although there was no reason to expect different patterns of results in the two cases. Different baselines for completion of fragments deriving from high- versus low-frequency words could have led to problems in interpreting whatever priming was observed, so we opted for a common baseline.

Achieving the common baseline could be accomplished by differentially deleting letters in a variety of ways. Although we did not systematically select one of these ways, the main difference between the fragment sets for the high- and low-frequency words would appear to be the extent of first-letter deletion. Fragments of high-frequency words were made harder to complete by deleting the first letter in 69% of the words; the corresponding value was 48% for fragments of low-frequency words. First letters are very influential in word fragment completion, so this difference probably was largely responsible for making the fragments of high-frequency words as hard to complete as those of low-frequency words.[1] We ensured, however, that the proportion of deleted letters was about the same for fragments of low-frequency words (.54) as for fragments of high-frequency words (.55).

Procedure. For the study phase, the program randomly selected 24 low-frequency words and 24 high-frequency words. These words were randomly ordered and presented in lowercase letters at the center of the monitor for 3 s each. Participants were instructed to read and study each word carefully because they would be asked to remember the words on a later memory test. The format of the memory test was not mentioned during the study phase.

---

[1] Tenpenny and Shoben (1992, Experiment 3) constructed their fragments so that they all contained the first letter of the word. This probably explains why their fragments were easier to solve than ours: Their average base rate of .43 was more than twice our .18 unprimed rate. Interestingly, although they did not intentionally equate baselines (see their footnote 2, p. 35), their baselines were in fact virtually identical. Their unprimed completion probabilities were .42 for fragments of low-frequency words and .44 for fragments of high-frequency words. For these values to be so close, some difference must have existed between the two sets of fragments that favored those of low frequency. Otherwise, the fragments of high-frequency words ought to have been easier to solve.

The test phase immediately followed the study phase. Two tests were administered: free recall and word fragment completion. Order of test administration was counterbalanced, with half of the participants receiving the tests in each of the two orders. Before performing the word fragment completion test, participants were told that this was part of a pilot study for another experiment and unrelated to the memory experiment. The intention was to make the word fragment completion test as indirect as possible, reducing the likelihood that participants would adopt intentional retrieval strategies.

For the word fragment completion test, participants saw fragments of 12 studied high-frequency words, 12 studied low-frequency words, 12 unstudied high-frequency words, and 12 unstudied low-frequency words, plus the 48 filler word fragments. The fillers were included as an additional way to disguise the relation of the word fragment completion test to the studied words, again making it as indirect as possible. These 96 fragments were presented in a unique random order for each participant. Each fragment appeared in lowercase letters, with missing letters replaced by underscores. This phase was also under computer control, with fragments presented at the center of the screen for a maximum of 15 s. If the participant produced a response before the 15 s expired, he or she could then press the space bar to advance to the next fragment. If the participant had not responded by 15 s, a beep sounded to warn the participant that the next fragment was now present on the screen.

For the free-recall test, participants were given 5 min to write down as many of the 48 studied words as they could remember. After the second of the two tests, participants were questioned about what they had noticed on the word fragment completion test and were then debriefed. The entire experiment lasted approximately 25 min.

Results and Discussion

Free recall. Table 1 displays the data from this experiment. The free-recall data, consisting of proportions of words correctly recalled, were analyzed by using a 2 × 2 mixed analysis of variance (ANOVA). The within-subject variable was word frequency; the between-subject variable was order of test administration. The major result for recall was as predicted: There was no effect of word frequency ($F < 1$). As anticipated in a mixed-list design, low-frequency words (.16) were recalled just as well as high-frequency words (.15). The absence of a frequency effect in recall was not due to a lack of statistical power, however; the power was .84 to detect a frequency difference of 5% or greater.

Nor did word frequency interact with the order in which the tests were administered ($F < 1$). However, participants who did the recall test first (.18) recalled significantly more words than did participants who did the recall test second (.13), shown by the reliable effect of order of testing, $F(1, 22) = 4.56$, $MSE = 0.007$, $p < .05$. Apparently, insertion of the word fragment completion test before the recall test resulted in some forgetting, possibly because of interference.

Word fragment completion. The word fragment completion data, consisting of proportions of fragments correctly completed, were analyzed by using a 2 × 2 × 2 mixed ANOVA, the additional within-subject variable being whether a given fragment was studied or not studied in the first phase. As anticipated, there was a significant effect of word frequency, $F(1, 22) = 6.81$, $MSE = 0.020$, $p < .05$. The fragments of low-frequency words (.32) were completed about 7% more often than those of the high-frequency words (.25).

Not surprisingly, the main effect of whether a word had previously been studied (.39) or not studied (.17) was also significant, $F(1, 22) = 45.56$, $MSE = 0.024$, $p < .001$, a basic priming effect of about 22%. However, the main effect of order of test administration was nonsignificant, $F(1, 22) = 1.41$, $MSE = 0.044$, $p > .20$, in keeping with the previously observed absence of retention interval effects on word fragment completion (e.g., Tulving, Schacter, & Stark, 1982).

With one crucial exception, all of the interactions were nonsignificant (all Fs < 1). The one exception was that word frequency interacted with whether a word had been studied in the first phase, $F(1, 22) = 15.23$, $MSE = 0.016$, $p < .001$. For words that had not been studied, there was no frequency effect, with fragments of high-frequency words (.18) completed as often as their low-frequency counterparts (.16). (Note that both of these values are close to the .20 baseline intended on the basis of the selection criteria for items from the pilot study.) For words that had been studied, fragments of low-frequency words (.47) were considerably more likely to be completed than those of high-frequency words (.29). Thus, the frequency effect was confined to the studied (primed) words.

These data displayed a dissociation between a direct recall test and an indirect fragment completion test with respect to the variable of word frequency. Although word frequency did not affect free recall, it did affect word fragment completion, with low-frequency words benefiting more from priming than high-frequency words. This confirmed the pattern reported by MacLeod (1989b) and by Roediger et al. (1992, Experiment 3) and also showed that the low-frequency advantage in word fragment completion priming can be large (22% in this experiment). The result is, however, in sharp contrast to the opposite pattern reported by Tenpenny and Shoben (1992, Experiment 3), which is the focus of Experiment 3.

Table 1
Experiment 1: Proportion of Words Correctly Recalled as a Function of Word Frequency and Order of Testing and Proportion of Word Fragments Correctly Completed as a Function of Word Frequency, Order of Testing, and Whether the Word Was Studied

	Recall first		Fragment completion first	
Condition	Low frequency	High frequency	Low frequency	High frequency
	Recall			
Studied				
M	.18	.18	.14	.12
SD	(.10)	(.09)	(.07)	(.07)
	Fragment completion			
Studied				
M	.44	.27	.51	.32
SD	(.25)	(.14)	(.20)	(.21)
Unstudied				
M	.15	.15	.17	.22
SD	(.10)	(.12)	(.10)	(.10)
Priming				
M	.29	.12	.34	.10
SD	(.26)	(.10)	(.16)	(.24)

Experiment 2

In the second experiment, using item recognition and word fragment completion tests, we predicted that word frequency would affect both types of test in the same way. For recognition, this expectation was based on the accumulated literature; for word fragment completion, it derived from the findings of Experiment 1. Experiment 1 demonstrated better word fragment completion of fragments derived from low-frequency words than of those derived from high-frequency words, and we expected to replicate this in Experiment 2. It has also been well established by researchers that recognition shows the same advantage for low-frequency over high-frequency words (Balota & Neely, 1980; Gorman, 1961; Shepard, 1967). Thus, there should be no dissociation in a study involving a direct recognition test and an indirect word fragment completion test.

Method

Participants. The participants were 24 volunteers from introductory psychology classes at the Scarborough Campus of the University of Toronto who received bonus points toward their grade for taking part in the experiment. None had taken part in Experiment 1.[2]

Materials and apparatus. These were exactly the same as in Experiment 1.

Procedure. The study phase was conducted exactly as in Experiment 1. Once again, the order of the two tests was counterbalanced, with half of the participants receiving them in each of the two orders. Each of the tests contained 48 critical items such that there was no overlap between the two tests. Thus, each test contained a different set of 12 studied low-frequency words, 12 studied high-frequency words, 12 unstudied low-frequency words, and 12 unstudied high-frequency words. This arrangement prevented any priming between the items on the successive tests.

The word fragment completion test was carried out exactly as in Experiment 1, again including the 48 filler items to disguise its relation to the study phase. The recognition test was also administered under computer control, with participants instructed to press a designated yes or no key as quickly as possible to indicate whether the particular word appearing at the center of the screen had appeared on the study list. The next test word immediately followed the participant's response. The entire experiment took about 25 min to carry out.

Results and Discussion

Recognition. Table 2 presents the data for the recognition and word fragment completion tests. Consider first the hit rate in recognition. A 2 × 2 mixed ANOVA was conducted on these data. As predicted, word frequency affected recognition hits significantly, $F(1, 22) = 5.04$, $MSE = 0.029$, $p < .05$; low-frequency words (.80) were better recognized than high-frequency words (.69). Word frequency did not interact significantly with testing order ($F < 1$), but there was a main effect of testing order, $F(1, 22) = 14.86$, $MSE = 0.031$, $p < .001$. Recognition hit rate was about 20% higher when the recognition test preceded (.84) than when it followed (.64) the word fragment completion test. As with recall in Experiment 1, and consistent with usual findings on direct tests, there was evidence of forgetting on the recognition test because of the intervening word fragment completion test.

Table 2
Experiment 2: Proportion for Hit Rate and False-Alarm Rate, and Hit Latency (in Milliseconds) in Recognition as a Function of Word Frequency and Order of Testing and Proportion of Word Fragments Correctly Completed as a Function of Word Frequency, Order of Testing, and Whether the Word Was Studied

Condition	Recognition first: Low frequency	Recognition first: High frequency	Fragment completion first: Low frequency	Fragment completion first: High frequency
Recognition				
Hit rate				
M	.90	.78	.69	.59
SD	(.11)	(.15)	(.16)	(.25)
False-alarm rate				
M	.09	.12	.16	.28
SD	(.12)	(.12)	(.15)	(.15)
Hit latency				
M	836	843	987	1251
SD	(166)	(188)	(187)	(378)
Fragment completion				
Studied				
M	.53	.29	.53	.35
SD	(.20)	(.17)	(.17)	(.17)
Unstudied				
M	.19	.18	.14	.24
SD	(.20)	(.14)	(.12)	(.14)
Priming				
M	.34	.11	.39	.11
SD	(.24)	(.20)	(.16)	(.14)

The false-alarm data converged on the same conclusions. Word frequency affected false-alarm rate, $F(1, 22) = 4.39$, $MSE = 0.016$, $p < .05$, with about 7% more false alarms for high-frequency words (.20) than for low-frequency words (.13). Glanzer and Bowles (1976) and Clark and Burchett (1994) reported this same logical reversal from hits to false alarms. Word frequency did not interact with testing order, $F(1, 22) = 1.76$, $p = .20$, but there was a significant main effect of testing order, $F(1, 22) = 8.96$, $MSE = 0.018$, $p < .01$. Not surprisingly, there were more false alarms—about 12% more—over all items when recognition was tested after (.22) than when it was tested before (.10) word fragment completion.

The recognition latency data converged on the same pattern as the accuracy data. A 2 × 2 ANOVA confirmed a significant main effect of frequency, $F(1, 22) = 5.44$, $MSE = 40{,}667$, $p < .05$, with low-frequency words (911 ms) recognized faster than high-frequency words (1,047 ms). Overall, words were also recognized faster if the recognition test was done first (839 ms) than if it was done second (1,119 ms), $F(1, 22) = 11.79$, $MSE = 79{,}661$, $p < .01$. These results are quite in keeping with the

[2] The initial group of 12 participants who did fragment completion first contained participants whose first language was not English and who were noted at the time of participation as having great difficulty in doing the word fragment completion task. Because of their poor scores and the resulting extremely high variance, this group was discarded, and an entirely new replacement group of 12 participants was tested.

accuracy data. The only unanticipated result in the latency data was the significant interaction of word frequency and testing order, $F(1, 22) = 4.85, p < .05$, showing that the latency advantage for low-frequency words occurred only when recognition was tested after word fragment completion. Given that the sets of items on the two tests were completely nonoverlapping, we have no ready interpretation of this peculiarity that did not appear in the accuracy data.

Word fragment completion. The word fragment completion data provided a thorough replication of the results in Experiment 1, as can be seen by comparing Table 2 with Table 1. The analysis was the same 2 × 2 × 2 mixed ANOVA used in Experiment 1. Once again, more fragments of low-frequency words (.35) were completed than of high-frequency words (.27), a reliable 8% difference, $F(1, 22) = 5.91$, $MSE = 0.028$, $p < .05$. Order of testing had no impact on performance ($F < 1$), but whether a word had been studied (.43) or not studied (.19) resulted in the anticipated large priming effect, $F(1, 22) = 50.93$, $MSE = 0.026$, $p < .001$. Note that the magnitude of priming in Experiment 2 (about 24%) was very close to that observed in Experiment 1 (about 22%).

As in Experiment 1, there was a significant interaction between word frequency and whether a word had been studied, $F(1, 22) = 42.43$, $MSE = 0.009$, $p < .001$. Previously unstudied words showed no frequency effect, with fragments of high-frequency words (.21) actually completed slightly more often than those of low-frequency words (.16); both values again were close to the intended baseline of .20. However, for studied words, fragments of low-frequency words (.53) were much more likely to be completed than were fragments of high-frequency words (.32). The frequency effect was confined to studied words. None of the remaining interactions were significant, the only F value exceeding unity being that for the Testing Order × Word Frequency interaction, $F(1, 22) = 1.86$, $p > .10$.

These data displayed a continuity between a direct recognition test and an indirect word fragment completion test; there was no suggestion of any dissociation. Both the recognition and word fragment completion tests showed a frequency effect in which performance was better for low-frequency words than for high-frequency words.

Experiment 3

Experiments 1 and 2 presented a clear picture. Priming in word fragment completion showed an advantage for low-frequency words, consistent with the post hoc analysis in MacLeod (1989b) and with the finding of Roediger et al. (1992, Experiment 3), although the present experiments showed that this low-frequency advantage can be quite large. Why, then, did Tenpenny and Shoben (1992, Experiment 3) observe the reverse—a small priming advantage for high-frequency words—on the same task? Experiment 3 was conducted to address this question directly. For this experiment, the materials used by Tenpenny and Shoben were substituted for the materials used in the prior experiments.

There were a number of differences between our first two experiments and the Tenpenny and Shoben (1992) experiment. We consider six of these. First, it could be that their use of capital letters as compared with our use of lowercase letters somehow affected performance, perhaps because of featural differences. We presented their materials in lowercase, in keeping with our previous procedure. Second, although we gave 15 s per fragment during the test, they gave only 10 s. We reduced our time to 10 s in Experiment 3. Third, whereas we equated baseline completion probability by differential letter deletion (particularly of initial letters) in high- versus low-frequency words, they did not, although they managed to obtain equal baselines nonetheless. Interestingly, their overall baseline level (.43) was much higher than ours (.18). Note, though, that Roediger et al. (1992) had somewhat higher baselines for high-frequency words, and an intermediate mean baseline value of .25, and that they did not use differential first-letter deletion. Therefore, the baseline issue does not appear to be a major concern. We suspect that these factors were not the critical ones.

We saw the remaining three differences between our Experiments 1 and 2 and Tenpenny and Shoben's Experiment 3 (1992) as most likely to be the ones that mattered. First, for reasons related to the main thrust of their article, Tenpenny and Shoben had participants study word pairs that were made up of a graphemically related cue word and a target word (e.g., *BRACKET–BRACELET*). Although the cues did not reappear at the time of test, this graphemic emphasis at study may have influenced how participants encoded the target words. We tested two groups of participants, one with these cues at study and one without the cues, which corresponded to their procedure and ours. We expected to replicate our pattern with single words but thought it possible that we would get their pattern when cues were studied together with the critical words.

The final two differences relate to how well the purpose of the word fragment completion test was disguised. In our experiments, considerable effort was exerted in this regard. Participants were told that the test was an unrelated pilot experiment, and no mention was made of the relation to study. Furthermore, partly because of the filler items, the proportion of studied words on the test was low: .25 in Experiment 1 and .33 in Experiment 2. In contrast, Tenpenny and Shoben (1992) explicitly informed their participants of the study–test relation: "Fragment-completion subjects were informed that some of the fragments were fragments of target words and others were fragments of words they had not studied" (p. 31). In addition, the proportion of studied items on the test was an unusually high .67. Taken together, these two differences seemed to us quite likely to have led to intentional retrieval of the studied items by participants in the Tenpenny and Shoben experiment, thereby essentially rendering an indirect test direct. We did not tell the participants in our final experiment about the study–test relation, and we added filler items to the test to reduce the proportion of studied items on the test to .33.

Put simply, we melded our design with their items. Every effort was made to disguise the nature of the word fragment completion test, making it as indirect as possible. Our design also permitted us to compare two groups: one that had word pairs at study and one that had single words. There was one last change with regard to Tenpenny and Shoben (1992): For comparison purposes with our Experiment 1, we added a free-recall test after the word fragment completion test. Finally, several smaller concerns were addressed, including font, solution time, and baseline issues. In summary, the focus of this experiment was on whether we would obtain our frequency effect (low-frequency words produce more priming) or theirs (high-frequency words produce more priming) on the indirect word fragment completion test and on whether the effect would differ as a consequence of studying critical words alone or in graphemic pairs.

Method

Participants. The participants were 56 students from introductory psychology classes at the Scarborough Campus of the University of Toronto who received bonus points toward their grade for taking part in the experiment. Twenty-eight were assigned to each of the two groups (which are described in the *Procedure* section). None had taken part in either Experiments 1 or 2. The data of 3 other participants who initially took part in this experiment were discarded without examination because their English was not sufficient to accommodate the low-frequency words.

Materials and apparatus. The critical high- and low-frequency words were those used by Tenpenny and Shoben (1992, Experiment 3, Appendix C). There were 36 high-frequency words, with a mean Kučera and Francis (1967) frequency of 150.11 occurrences per million, and 36 low-frequency words, with a mean Kučera and Francis frequency of 2.28 occurrences per million. Each critical word (e.g., *bracelet*) had a corresponding graphemically related word (e.g., *bracket*) and a corresponding word fragment (e.g., b-ce-t). All word fragments included the first letter of the critical word and had unique solutions (except for the occasional very rare alternative). With respect to their corresponding critical word, all graphemic cues also had the same first letter, were no more than two letters different in length, and contained no more than three different letters. Graphemic cues had to be semantically unrelated to their corresponding critical words as well. All of the graphemic cues we used were from the "highly similar" subset used by Tenpenny and Shoben (1992).

In addition, there were 48 filler words and corresponding fragments, consisting of the same fillers as used in Experiments 1 and 2, which are shown in the Appendix. The apparatus was identical to that of Experiments 1 and 2, except that a color VGA monitor replaced the monochrome monitor, and all displays were white letters on a black background.

Procedure. The two groups of participants differed only in the study phase of the experiment. The *one-word* group was treated very similarly to the participants in Experiments 1 and 2. They studied 36 words—18 low frequency and 18 high frequency—presented one at a time in lowercase at the center of the screen. For each participant, the studied words were randomly selected from the entire set and presented in a unique random order. The study phase began with a 1-s warning row of asterisks. Each word was then presented for 3 s, followed by a 250-ms blank screen. Another 1-s row of asterisks followed the last study word.

The *two-word* group saw two words centered on the screen on each study trial. The two words were always a graphemic cue word in uppercase on the left and its corresponding critical word in lowercase on the right, separated by a hyphen (e.g., *BRACKET–bracelet*). This study procedure mimicked that of Tenpenny and Shoben (1992). Participants were told to learn the lowercase word on the right by using the uppercase word on the left to assist them.

Before study, both groups were told the length of the study list and were told to remember the words for a later memory test. The indirect test phase was identical for the two groups. The instructions began as follows: "Before the memory test, we would like you to do another task. As you may have guessed, this is to put some time between study and test, making it harder for you to remember later." Of course, this was the word fragment completion test, which was referred to throughout the instructions as the *intervening task* to deemphasize its relation to the study list. Participants were encouraged to try hard but were told that the fragments were difficult for everyone and they should not get discouraged if they missed quite a few. The participants were asked not to use proper nouns as solutions.

In keeping with Tenpenny and Shoben's procedure (1992), test time for each fragment was reduced from the 15 s in our Experiments 1 and 2 to the 10 s that they had used. Participants were to respond aloud and were cautioned that only their first response to each fragment could be accepted. If they did respond within 10 s, the experimenter input a keypress to indicate whether they had produced the right word or a wrong word. Participants were told that the computer would go on to the next fragment after 10 s if they had not responded to the current fragment. A beep coincided with the onset of each new fragment to ensure that participants did not miss any. There was a 250-ms blank between successive fragments.

The test sequence consisted of a random ordering of the 36 fragments of low-frequency words (18 studied and 18 unstudied), the 36 fragments of high-frequency words (18 studied and 18 unstudied), and 36 filler fragments, randomly selected from the 48 available. Thus, there were 108 fragments in all. The experimenter had a protocol available so that she could determine for each item whether the participant had responded correctly; if not, she could then write down the participant's incorrect response.

The direct free-recall test was very basic. After finishing the indirect word fragment completion test, participants were asked to try to write down as many words as they could from the originally studied list. Participants in the two-word condition were also told to try especially to recall the second (critical) word of each pair but to write down any word, even if they were unsure. No time limit was assigned for recall.

Results and Discussion

Table 3 presents the free-recall and word fragment completion data separately for words studied with versus without a graphemic cue word. In this experiment, because the main question related to priming on the word fragment completion test, that test was always administered before the recall test, so there is no order of testing variable reported for Experiment 3.

Recall. The recall analysis was a 2 × 2 mixed ANOVA, the between-subjects variable being number of words studied (cued was two words and uncued was one word) and the within-subject variable being word frequency. As in Experiment 1, there was no effect of word frequency ($F < 1$), despite a power of .99 to detect a 5% frequency difference in recall. As Table 3 shows, low-frequency words (.12) were recalled as often as high-frequency words (.12). Furthermore, the main effect of number of words at study was not reliable ($F < 1$), nor was the interaction of this variable with word frequency, $F(1, 54) = 1.89$, $MSE = 0.002$, $p = .18$. Neither the cuing manipulation nor the word frequency manipulation affected recall, the latter being consistent with the finding of Experiment 1.

Word fragment completion. A 2 × 2 × 2 mixed ANOVA was carried out on the word fragment completion data, with the between-subject variable being number of words at study

Table 3
Experiment 3: Proportion of Words Correctly Recalled as a Function of Word Frequency and Whether the Word Was Studied With or Without a Graphemic Cue and Proportion of Word Fragments Correctly Completed as a Function of Word Frequency, Whether the Word Was Studied, and Whether the Word Was Studied With or Without a Graphemic Cue

	Without cue		With cue	
Condition	Low frequency	High frequency	Low frequency	High frequency
	Recall			
Studied				
M	.13	.11	.11	.13
SD	(.06)	(.06)	(.06)	(.07)
	Word fragment completion			
Studied				
M	.63	.66	.60	.66
SD	(.13)	(.10)	(.12)	(.13)
Unstudied				
M	.34	.52	.41	.53
SD	(.15)	(.13)	(.13)	(.13)
Priming				
M	.29	.14	.19	.13
SD	(.18)	(.15)	(.14)	(.19)

and the within-subject variables being word frequency and whether the word had been studied.

There was no overall difference in performance between the one-word and two-word groups ($F < 1$). Number of studied words did not interact with word frequency alone ($F < 1$) or in the three-way interaction with word frequency and whether the word was studied, $F(1, 54) = 1.90$, $MSE = 0.015$, $p = .17$. However, the interaction of number of words at study with whether a word was studied, $F(1, 54) = 3.13$, $MSE = 0.013$, $p = .08$, was marginally significant. This can be seen in Table 3 and reflects the fact that there was less priming for critical words studied in pairs (.16) than for critical words studied alone (.22). This is not an unprecedented result: MacLeod (1989b) showed that embedding words in context at study diminished the extent of priming they caused on a word fragment completion test. Critically, though, in this experiment, this context effect was independent of word frequency.

The key results were virtually identical to those reported in Experiments 1 and 2, although the overall magnitude of priming declined. As before, there was a reliable main effect of word frequency, $F(1, 54) = 48.81$, $MSE = 0.011$, $p < .001$, with fragments of high-frequency words (.59) completed about 9% more often than those of low-frequency words (.50). There was also a significant main effect of whether a word had been studied, $F(1, 54) = 153.28$, $MSE = 0.013$, $p < .001$, with studied words (.64) completed more often than unstudied words (.45). Most important, these two variables interacted reliably, $F(1, 54) = 9.83$, $MSE = 0.015$, $p < .01$. Averaged over the two groups, the magnitude of priming was about 10% greater for low-frequency words (.24) than for high-frequency words (.14).

Quite clearly, Experiment 3 produced the same pattern as in Experiments 1 and 2. Priming in word fragment completion was greater for low-frequency words than for high frequency words, although by about 10% here as compared with the 22% earlier.[3] The number of words at study had almost no effect, with only a suggestion that priming for low-frequency words was somewhat reduced when the words were studied with graphemically related pairs.

It would appear that the standard result is greater priming for low-frequency words in word fragment completion. Our strong surmise is that Tenpenny and Shoben (1992) observed the opposite result because participants in their experiment were using an intentional retrieval strategy. Knowing that most of the test fragments had been presented during study, their participants consciously tried to recover studied items, thereby contaminating the nominally indirect word fragment completion test and, in essence, turning it into a direct test. Although we have not directly tested this intentional retrieval hypothesis, it seems very likely to be the reason for the discrepant results of Tenpenny and Shoben.

General Discussion

In summary, word frequency affected both item recognition and priming in word fragment completion in the same way: Performance was better on low-frequency words than on high-frequency words. There was no effect of word frequency on free recall. The major finding is the low-frequency word advantage in word fragment completion, which was replicated in our three experiments with two distinct sets of materials. The recall and recognition results, of course, confirmed findings already described in the literature. The critical finding is that word fragment completion, an indirect test of memory, showed the same pattern as one direct test, recognition, but not the other, recall. Put another way, performance on the two direct tests was dissociated, but there was no dissociation between one direct and one indirect test.

The frequency effect in word fragment completion was consistent with the repetition priming patterns observed on other indirect tests of memory such as masked word identification (Jacoby & Dallas, 1981) and lexical decision (Scarborough et al., 1977). The fact that the same pattern was evident in recognition suggests at least some processing overlap between these two classes of memory tests. Like certain other manipulations (e.g., directed forgetting in MacLeod, 1989a), word frequency influences at least some direct and indirect tests in the same way. Indeed, these two experiments showed that both parallel and dissociative effects can be observed across direct and indirect tests.

[3] Interestingly, the baselines did differ in Experiment 3: For unprimed items, participants completed fragments of high-frequency words about 14% more often than fragments of low-frequency words. This would be expected, given that baselines were free to vary, and is consistent with the pattern reported by Roediger et al. (1992, Experiment 3) who observed about an 8% difference in their free-to-vary baselines. It remains puzzling, though, why the baselines did not differ in the Tenpenny and Shoben study, in which they also were not intentionally equated.

Of course, it is possible that intentional recollection is intruding into the nominally implicit performance of the word fragment completion task. Participants may notice some studied words and begin conscious attempts to recover words from the study list. This is a potential problem with all implicit tests; Richardson-Klavehn and Bjork (1988, p. 528) pointed out that "parallel effects will remain difficult to interpret (*without criticisms*) until we know more about the contribution of intentional memory to performance on implicit tests." However, this criticism is more difficult to sustain in the face of two aspects of our results. First, the pattern differed for the two direct tests, in which intentional memory was necessarily involved in both cases. Second, it is very likely that Tenpenny and Shoben's (1992) results in their Experiment 3 were contaminated by intentional recollection, a problem less likely to have occurred in our Experiment 3. Given the opposite findings of these two experiments, this makes an explanation of our results in terms of explicit intrusion very difficult.

A plausible explanation of our results, and many other results, is provided by Mandler's (1980) two-process recognition theory. This theory has the virtue of having already been applied to explain the word frequency effect (Mandler, 1980) and to account for performance on direct versus indirect tests of memory (e.g., Graf & Mandler, 1984; Graf & Ryan, 1990; Mandler, 1989). In the theory, Mandler discriminates between familiarity and retrieval as bases for recognition. A given item on a recognition test may cause a strong sense of familiarity, in which case the participant will respond positively. Alternatively, if the level of familiarity is insufficient to reach the threshold for responding, the participant may then attempt a retrieval. If that retrieval succeeds, the participant will make a positive response. Thus, recognition can occur in either of two ways. However, the familiarity route is tried first and, hence, often dominates.

Mandler (1980) argued that familiarity develops through progressive integration–organization within an item over repeated encounters. Familiarity is little influenced by connections between items. Naturally, low-frequency words have had fewer opportunities to increment familiarity. Thus, when a low-frequency word is presented, the enhanced integration and the consequent increment to its familiarity are much greater than is the case when a high-frequency word is presented. Recognition tests are very sensitive to this increment in familiarity (e.g., Hintzman, 1976), and, hence, low-frequency words are better recognized than high-frequency words.[4]

The second process is retrieval, and the probability of retrieval increases to the extent that there has been elaborative processing of an item that has gone beyond the item itself, linking it to its context and to other information already in memory. High-frequency words will tend to be elaborated more than low-frequency words, and hence high-frequency words will tend to be more retrievable. This is an advantage primarily on a recall test, in which the items themselves are not presented again and the participant must rely on self-initiated retrieval: The consequence is a high-frequency superiority in recall. This superiority may apply only to pure lists because in such lists the ease of connecting high-frequency words to each other is at its greatest, without the disruption of hard-to-integrate low-frequency words. In line with this argument, Mandler (1980, p. 268) suggested that in a mixed-list, low-frequency words would receive "at least some additional attention" relative to high-frequency words. This diversion of attention could help to minimize the frequency effect both by disrupting elaboration of high-frequency words and by improving the encoding of low-frequency words.

We must assume, following Graf and Ryan (1990), that study episodes lead to both integrative and elaborative processing, the amount of each depending on the conditions of study. In the present experiments, we did not manipulate study conditions (apart from the quite inconsequential context manipulation of Experiment 3), so this assumption is not especially relevant here. So the key assumption for us is that word fragment completion, like recognition, relies more on integrative processing and familiarity than does recall, which emphasizes elaborative processing and retrieval. Indeed, we would go a step farther, endorsing Graf and Ryan's (1990) suggestion that most indirect tests of memory may be especially sensitive to integrative processing that increases familiarity.

Under this view, then, recognition and priming in word fragment completion should behave similarly. Both should benefit from the greater boost in familiarity accruing from the presentation of a low-frequency word as compared with a high-frequency word during study. Both tests are sensitive to changes in familiarity. Recall, dependent as it is on interitem elaborative processing, should show little word frequency effect if linking the high-frequency words is made difficult and attention to the low-frequency words is promoted, as when low-frequency words are mixed with the high-frequency words. Because we used mixed frequency lists, we therefore did not observe a high-frequency advantage in recall.

This interpretation places recognition, a direct test, much closer to indirect tests, such as word fragment completion, in terms of the fundamental processes tapped by these types of tests. To the extent that we have observed a dissociation among tests, then, it is between two tests that rely on the familiarity gain that accrues from integration—recognition and word fragment completion—and one test that relies on strengthening by elaboration—recall. The dissociation is not between direct and indirect tests but between the different processes that underlie different tests. Thus, we strongly agree with Witherspoon and Moscovitch (1989) who, having found a dissociation between two indirect tests, argued that "the degree of dependence between performances on memory tests is determined by the similarity of the component processes that the tests engage and of the information they use" (p. 22). The word frequency effect provides an excellent illustration of this principle.

[4] Interestingly, Gardiner and Java (1990) reported that word frequency exerted its influence in recognition on *remember* responses, in which participants had a recollective experience, but not on *know* responses, which presumably reflect only feelings of familiarity.

References

Balota, D. A., & Neely, J. H. (1980). Test-expectancy and word-frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 576–587.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition: An analysis of semantic strategies. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 252–259.

Bousfield, W. A., & Cohen, B. H. (1955). The occurrence of clustering in the recall of randomly arranged words of different frequencies of usage. *Journal of General Psychology, 52,* 83–95.

Cattell, J. M. (1886a). The time it takes to see and name objects. *Mind, 11,* 63–65.

Cattell, J. M. (1886b). The time taken up by cerebral operations. *Mind, 11,* 220–242, 377–392, 524–538.

Clark, S. E. (1992). Word frequency effects in associative and item recognition. *Memory & Cognition, 20,* 231–243.

Clark, S. E., & Burchett, R. E. R. (1994). Word frequency and list composition effects in associative recognition and recall. *Memory & Cognition, 22,* 55–62.

Dahl, H. (1979). *Word frequency of spoken American English.* Essex, CT: Verbatim.

Duchek, J. M., & Neely, J. H. (1989). A dissociative word-frequency × levels-of-processing interaction in episodic recognition and lexical decision tasks. *Memory & Cognition, 17,* 148–162.

Duncan, C. P. (1974). Retrieval of low-frequency words from mixed lists. *Bulletin of the Psychonomic Society, 4,* 137–138.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18,* 23–30.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Glanzer, M., & Bowles, N. (1976). Analysis of the word frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 21–31.

Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology, 61,* 23–29.

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., & Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 978–992.

Graves, R., & Bradley, R. (1987). Millisecond interval timer and auditory reaction time programs for the IBM PC. *Behavior Research Methods, Instruments, & Computers, 19,* 30–35.

Graves, R., & Bradley, R. (1988). More on millisecond timing and tachistoscope applications for the IBM PC. *Behavior Research Methods, Instruments, & Computers, 20,* 408–412.

Gregg, V. (1976). Word frequency, recognition, and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183–216). New York: Wiley.

Gregg, V. H., Montgomery, D. C., & Castano, D. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior, 19,* 240–245.

Hall, J. F. (1954). Learning as a function of word frequency. *American Journal of Psychology, 67,* 138–140.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 47–91). New York: Academic Press.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word probability. *Journal of Experimental Psychology, 41,* 401–410.

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

Kintsch, W. (1970). *Learning, memory, and conceptual processes.* New York: Wiley.

Kirsner, K., Milech, D., & Standen, P. (1983). Common and modality-specific processes in the mental lexicon. *Memory & Cognition, 11,* 621–630.

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

MacLeod, C. M. (1989a). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 13–21.

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

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

Mandler, G. (1989). Memory: Conscious and unconscious. In P. R. Solomon, G. R. Goethals, C. M. Kelley, & B. R. Stephens (Eds.), *Memory: Interdisciplinary approaches* (pp. 84–106). New York: Springer-Verlag.

May, R. B., & Tryk, H. E. (1970). Word sequence, word frequency, and free recall. *Canadian Journal of Psychology, 24,* 299–304.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43–71.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76,* 340–354.

Peters, H. N. (1936). The relationship between familiarity of words and their memory value. *American Journal of Psychology, 48,* 572–584.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 475–543.

Roediger, H. L., III, Weldon, M. S., Stadler, M. L., & Riegler, G. L. (1992). Direct comparisons of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1251–1269.

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.

Schwartz, F., & Rouse, R. D. (1961). The activation and recovery of associations. *Psychological Issues, 3* (Whole No. 1).

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6,* 156–163.

Solomon, R. L., & Postman, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology, 43,* 195–201.

Sumby, W. H. (1963). Word frequency and the serial position effect. *Journal of Verbal Learning and Verbal Behavior, 1,* 443–450.

Tenpenny, 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., 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.

Underwood, B. J., Ekstrand, B. R., & Keppel, G. (1965). An analysis of intralist similarity in verbal learning with experiments on conceptual similarity. *Journal of Verbal Learning and Verbal Behavior, 4,* 447–462.

Wallace, W. P., Sawyer, T. J., & Robertson, L. C. (1978). Distractors in recall, distraction-free recognition, and the word frequency effect. *American Journal of Psychology, 91,* 295–304.

Witherspoon, D., & Moscovitch, M. (1989). Stochastic independence between two implicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 22–30.

Appendix

The 48 Low-Frequency Words, 48 High-Frequency Words, and 48 Filler Words from Experiments 1 and 2

Low frequency		High frequency		Filler	
adventure	mustang	analysis	medicine	amplifier	important
charade	navigator	article	mistake	answer	jealous
courier	nostalgia	average	morning	assassin	marriage
derelict	obscenity	birthday	negative	bachelor	network
dynamite	orchestra	careful	number	bedroom	normal
elephant	outcast	comment	original	blizzard	office
epidemic	pacifier	difficult	people	bouquet	opinion
festival	parachute	distance	percent	business	optimist
fiasco	polaroid	emotion	physical	chimney	passport
flannel	rebuttal	example	pleasant	cinnamon	pilgrim
habitual	riviera	fantasy	private	coconut	reclusive
heckler	schizoid	freedom	question	decision	religion
histamine	scorpion	friend	rabbit	direction	republic
hypnotist	sergeant	graduate	regular	doldrums	smuggler
insanity	strategic	ground	serious	eclipse	success
intestine	travesty	health	special	element	threshold
jackpot	trilogy	history	strength	euphoria	toboggan
jamboree	ultimatum	interest	surgery	forgery	utopia
legislate	villain	judgment	tonight	gangster	vacation
lipstick	voucher	junior	traffic	garage	volleyball
literary	wardrobe	kitchen	trouble	geography	weather
macaroon	warrantee	knowledge	various	husband	wonderful
marauder	watermelon	language	violent	illusion	yoghurt
mosaic	zoology	lecture	window	immunity	zenith

Note. The underlined letters in each word were replaced with an underscore on the test to create the corresponding word fragment.

Received August 17, 1994
Revision received December 23, 1994
Accepted January 19, 1995 ■