Altering memory through recall: The effects of cue-guided retrieval processing

MARK A. McDANIEL
Purdue University, West Lafayette, Indiana

and

MICHAEL D. KOWITZ and PAUL K. DUNAY
University of Notre Dame, Notre Dame, Indiana

Three experiments were designed to investigate the influence of initial recall on memory by assessing delayed recall after different immediate cued-recall tests. In all experiments, subjects performed semantic and phonemic encoding tasks on a word list. The subjects then received a cued-recall test that cued the target using the same word as the context word in the encoding task, a test that cued the target with a word from the same level at which the target was encoded, a test that cued the target with a cue from a different level at which the target was encoded, or no immediate-recall test. One day later, the subjects performed a final cued-recall test in which the type of cue (semantic or phonemic) was varied. Consistently, delayed recall was facilitated primarily when the cue on the immediate test was from the same level as the cue on the delayed test. This pattern of facilitation suggests that immediate cued-recall produces an elaboration of an existing memory representation that is closely tied to the type of cue used on the immediate test.

It is well documented that recall attempts of previously encoded targets can improve performance on later memory tests for those targets (e.g., Darley & Murdock, 1971; Hanawalt & Tarr, 1961; Hogan & Kintsch, 1971; Wenger, Thompson, & Bartling, 1980). This finding implies that once an episodic memory representation of an event is established, subsequent retrieval will alter the memory of that event (Bjork, 1975). The literature cited above does not, however, clearly illuminate the mechanisms by which retrieval influences memory. In recent studies, investigators have attempted to determine how retrieval of an event may alter the original encoding of the event. Bartlett (1977) presented subjects with lists of six words each, and immediately after presentation of each list, recall was cued with either nonsemantic (orthographic or temporal) cues or semantic (category) cues. Initial retrieval with the semantic cues enhanced retention on a subsequent final free-recall test more than did initial retrieval with the nonsemantic cues (see Whitten, 1978, for a similar finding). This finding suggests that initial retrieval operations are guided by the retrieval cues provided (Craik, 1981), with semantic cues prompting

deep, semantic processing and nonsemantic cues prompting shallow, nonsemantic processing (Bartlett, 1977).

The notion that retrieval processing depends on the type of retrieval cue provided appears plausible and is appealing from the standpoint of being able to recruit extant memory models (e.g., levels of processing; context-reinstatement models, Runquist, 1983) to understand the mnemonic effects of retrieval; however, the evidence bearing on this notion is scarce and somewhat contradictory. For example, employing a paradigm similar to that used by Bartlett (1977), Bartlett and Tulving (1974) reported that there was *no* effect of initial cue type (temporal or semantic) on delayed free recall for all except the last two input positions of a six-item list (recency items). Bartlett and Tulving concluded that "one and the same retrieval mechanism was involved in the immediate recall of all of these words [preterminal items]" (p. 307), with this mechanism leading to a recoding of the information about the word in memory that was common to the different initial retrieval cues. To assess more directly the nature of the recoding prompted by different retrieval cues, McDaniel and Masson (1985) used a cued-recall test for final recall. Their results suggested that initial retrieval increased the variability of the encoding of the target word.

The present experiments were designed to gather more information on the above issue. Using a paradigm based on that developed by McDaniel and Masson (1985), we manipulated the initial level of encoding of the target items, the type of cue provided at final recall, and the compatibility of the immediate-recall cue both with the initial level of encoding and with the final-recall cue. As outlined below, these experiments were designed such that

competing predictions could be derived from the different hypotheses about the mechanism(s) underlying the mnemonic effects of retrieval.

EXPERIMENT 1

In this experiment, subjects were required to process a word list using an incidental orienting task that varied the level of processing (semantic or phonemic) within subjects. Next, the subjects either were given one of two types of immediate cued-recall tests or were given no immediate-recall test. For one group of subjects, each target was cued by a word that was related to the target at the *same* level as the initial orienting task (the same-level immediate test). For example, if the orienting task had been to judge the phonemic aspects of LARK, the cue would be a phonemically related word such as BARK. Similarly, if the orienting task had been semantic (a category-related judgment), the cue would be a semantically related word such as ROBIN. Another group of subjects received an immediate test in which the target was cued by a word that was from the *alternative* level as the initial orienting task (the different-level immediate test). For this group, if LARK had been phonemically processed originally, the cue would be ROBIN; if LARK had been semantically processed, the cue would be BARK. A third group of subjects was given no immediate-recall test. This group served as a baseline against which the immediate-test groups' performance could be compared.

Following the immediate-recall test, 24 h later, all subjects received another cued-recall test in which each target word was cued with either a semantically or a phonemically related word (the delayed-recall test). This variable was factorially combined with the type of cue used on the immediate test so that half of the delayed-test cues were of the same cue type as those used on the immediate test (compatible cue cell; i.e., targets cued with semantic cues on the immediate test were also cued with semantic cues on the delayed test, and the same match occurred for phonemic cues) and the other half of the delayed-test cues were of the "opposite" cue type of those used on the immediate test (incompatible cue cell; i.e., targets cued with semantic cues on the immediate test were cued with phonemic cues on the delayed test, and vice versa). None of the cues on the delayed test had been previously used on the earlier tests. A cued-recall test was used because this test is sensitive to the kind of attributes represented in a target item's encoding (Fisher & Craik, 1977; Runquist, 1982).

We reasoned that this experiment would be sensitive to the possible mnemonic mechanisms of retrieval outlined at the outset. Specifically, the compatibility variable (the compatibility of immediate- and delayed-test cues) provides an indication of the degree to which retrieval processing corresponds to the type of cue provided at initial recovery (i.e., on the immediate test). If recovery with semantic cues produces semantic processing and recovery with nonsemantic (phonemic) cues produces non-

semantic processing, then the products of such processing should be to add (or strengthen) semantic and nonsemantic components, respectively, to the memory trace (McDaniel, Friedman, & Bourne, 1978) and/or to increase the availability of particular retrieval routes (semantic or nonsemantic, respectively) to the core encodings (Whitten & Bjork, 1977). In either case (added components or more available retrieval routes), delayed recall should be enhanced, relative to that of the control group, primarily when the delayed-retrieval cues are compatible with the immediate-retrieval cues. This result should obtain because, in keeping with the encoding-specificity principle (Tulving & Thomson, 1973), the information (or retrieval routes) added through initial retrieval is most likely to be accessed by a compatible final (delayed) retrieval cue.

Alternatively, if *similar* processing is invoked by different types of retrieval cues, then the compatibility of immediate-test cues and delayed-test cues would have little impact on the extent to which delayed recall is enhanced by immediate retrieval. For instance, it may be that, irrespective of immediate-cue type, successful retrieval adds primarily semantic information to an item's memory representation (Runquist, 1982). If this is the case, then semantic delayed-test cues should best evoke memory after immediate retrieval, regardless of immediate-cue type (in the analyses reported below, this would be manifested in a higher order interaction between the compatibility variable and the other variables).

Method

Subjects. The subjects were 56 male and female students at Indiana University at South Bend who participated for extra credit in their psychology courses. Forty subjects received an immediate-recall test, with 18 subjects randomly assigned to the same-level test group and 22 subjects assigned to the different-level test group. Sixteen subjects received no immediate-recall test.

Materials. The stimuli were 40 unrelated concrete nouns taken from the 48 used by McDaniel and Masson (1985). Four words at each end of the list were designed as buffer words to control for primacy and recency effects, leaving 32 items as targets. For the encoding phase of the experiment, a pamphlet of questions was assembled, one question per stimulus item. The questions were of the form, "How well does the word categorize [rhyme] with [an appropriate word]?" For example, when shown the target word LARK, the question might be either, "How well does the word categorize with JAY?" or, "How well does the word rhyme with BARK?" Half of the questions involved categorization and half involved rhyming.

For the retrieval phases of the experiment, three cued-recall tests were constructed. Two cued-recall tests were employed for immediate recall: one test for the same-level immediate-test group and one test for the different-level immediate test group. For both tests, the cues consisted of either a word that was a member of the same category as the target item or a rhyme of the target item. The words used as cues were not the same as those used in the encoding questions. For the same-level test, all of the semantically encoded targets were cued with a category member and all of the phonemically encoded targets were cued with a rhyme. For the different-level test, all of the semantically encoded targets were cued with a rhyme and all of the phonemically encoded targets were cued with a category member. To the left of each cue, a "C" or an "R" was

typed to indicate the relationship of the cue to the target, and all of the cues were typed on a single response sheet.

The other cued-recall test was employed for delayed testing. The cues were either rhymes or category associates and were different from those used on the immediate-recall tests. For the delayed test, half of the targets cued semantically (i.e., with a category member) on the immediate test were cued with a category member, and the other half were cued with a rhyme. The same held true for those targets that were phonemically cued on the immediate test. Thus, the cues were assigned such that there was an equal number of target items (eight) in each of the four possible cells formed by the factorial combination of type of immediate cue and type of delayed cue. The assignment of target items to these four categories was counterbalanced across subjects.

For both the immediate and the delayed tests, an attempt was made to equate the phonemic and semantic cue for a particular target in terms of the facility with which the target could be generated from each cue. This was done by starting with a pool of three words that were phonemically related to each target (96 total) and three words that were semantically related to each target (96 total). Six sets of booklets were constructed, each containing a phonemic cue for half of the targets and a semantic cue for the other half of the targets. A code letter ("C" or "R") was printed beside each cue indicating the type of associate to be generated. Each cue for a particular target was used once across the six booklets. The booklets were randomly distributed to 81 students from a University of Notre Dame introductory psychology class (participating for extra credit), with the number of students receiving each booklet ranging from 10 to 17. The students were instructed to generate associates (of the type designated) to each cue word during successive 15-sec intervals. From these responses, the probabilities of generating the targets given a particular cue were calculated. Cues were then selected for each test such that the probability of generating a target did not significantly differ across the two cue types. For 13 targets, the available pool of cues did not provide a pair of cues (i.e., a category cue and a rhyme cue) that could be equated for at least one test (immediate or delayed). Consequently, a new pool of cues was generated for these 13 targets, and 12 additional students generated responses to these cues. From the enlarged cue pool, cues that were closest in terms of target-generation probabilities were selected to be used for the 13 targets (for which cues had not yet been equated).

Procedure. The subjects were randomly assigned to one of the three groups: an initial-retrieval group that received the different-level immediate test, an initial-retrieval group that received the same-level immediate test, and a control group that did not receive an immediate-recall test. Each subject was given the pamphlet containing the encoding questions and was instructed to answer each question by using a rating scale of 1 to 5, with 1 indicating "not very well" and 5 indicating "almost perfect." The subjects were warned that the target items were always associated to some degree with the word contained in the question, but they were to use any method they could to make discriminations in the degree of association. Before the word list was presented, the subjects were given a practice list of two words to familiarize them with the encoding task and presentation rate. The target words were presented for 8 sec each, using a slide projector. The subjects were not forewarned about the memory tests.

After the presentation list, depending on the group to which the subjects were assigned, they either were given the appropriate immediate cued-recall test or were dismissed. After 24 h, all subjects returned to the laboratory and were given the delayed cued-recall test. Both cued-recall tests were subject-paced. For both cued-recall tests, the subjects were instructed to "try to recall the target words by working from the cues ... note the relationship that a cue has with a target word and try to recall the target word on that basis." Finally, the subjects were asked whether they had expected to be given a memory test in the second session on the target items and,

if so, whether they had consciously rehearsed the test items during the retention interval. Three subjects indicated that they had anticipated the memory test and had rehearsed for it. Therefore, these 3 subjects were replaced.

Results and Discussion

Immediate recall. The mean proportions of words recalled on the immediate-recall tests are shown in Table 1. A two-factor mixed analysis of variance (ANOVA), with encoding level as a within-subjects factor and type of immediate test as a between-subjects factor, indicated that semantically encoded items were recalled significantly better than were phonemically encoded items [$F(1,38) = 74.43, MS_e = .01$] and that the test with same-level cues tended to produce higher recall than did the test with different-level cues [$F(1,38) = 4.04, MS_e = .03, p < .06$]. (Unless otherwise noted, the alpha level for all statistical tests equals .05.) The encoding level and immediate-test type effects did not interact ($F < 1$). These results parallel other findings showing combined effects of levels of processing and encoding specificity (e.g., Fisher & Craik, 1977; McDaniel et al., 1978).

Delayed recall. Table 2 gives the mean proportions of words recalled by the immediate-test groups and the no immediate-test control group (first four columns of the means). The recall scores were analyzed with a three-factor mixed ANOVA, with immediate-test experience (no immediate test, same-level immediate test, or different-level immediate test) as a between-subjects variable and encoding task and type of delayed cue as within-subjects variables. To avoid confusion with the subsequent analyses reported (and with the introduction), some preliminary comment on the analysis of the unconditional recall data is needed. For this ANOVA, the compatible immediate-cue/delayed-cue cells could not be meaningfully constructed for the control group because the particular control cells used to assess compatibility effects changed, depending on the immediate-test group of interest (see Table 2). Consequently, in this analysis, the

Table 1
Mean Proportion Recalled on the Immediate-Recall Tests in Experiments 1, 2, and 3

Encoding Task	Immediate Test Type		
	Same Level	Different Level	Identical Word
Experiment 1			
Semantic	.42	.33	
Phonemic	.18	.12	
Experiment 2			
Semantic	.59		.82
Phonemic	.38		.47
Experiment 3			
Semantic	.51	.54	
Phonemic	.35	.25	

Note—Same Level = cues relating to the target at the same level as original encoding but the cues were different words than those used in the encoding context, Different Level = cues relating to the target at a different level than original encoding, Identical Word = cues were the identical words used in the original encoding context.

Table 2
Mean Proportion Recalled on the Delayed Test, Mean *a* Scores, and Mean Difference Scores in Experiment 1

Encoding Task/ Immediate-Recall Cue	Same Level Immediate-Test Group		Control Group		*a* Scores and Difference Scores	
	Compatible	Incompatible	Compatible	Incompatible	Compatible	Incompatible
Semantic Encoding/ Semantic Cue	(s) .31	(p) .15	(s) .27	(p) .23	(s) .26 (.05)	(p) −.03 (−.06)
Phonemic Encoding/ Phonemic Cue	(p) .17	(s) .06	(p) .14	(s) .08	(p) .17 (.03)	(s) −.04 (−.01)
Mean	.24	.10	.20	.16	.29 (.04)	−.04 (−.04)
	Different-Level Immediate-Test Group					
Semantic Encoding/ Phonemic Cue	(p) .26	(s) .30	(p) .23	(s) .27	(p) .16 (.04)	(s) .24 (.03)
Phonemic Encoding/ Semantic Cue	(s) .12	(p) .14	(s) .08	(p) .14	(s) .48 (.06)	(p) .20 (0.0)
Mean	.19	.22	.16	.20	.32 (.04)	.22 (.02)

Note—Compatible = compatible immediate and delayed cues, Incompatible = incompatible immediate and delayed cues. (s) = semantic delayed cue, (p) = phonemic delayed cue. For the *a*-score calculation, a reduced set of subjects in the immediate-test conditions was used (see text for details); the tabled mean proportion recall values reflect performance of all subjects. The control group received no immediate test. The set of control scores is reported successively to indicate the appropriate control value for each cell of the immediate-test groups. The *a* scores are based on Lockhart's (1975) measure (see text for full explanation). Difference scores are in parentheses and are based on the difference between proportion recalled by subjects given an immediate test and proportion recalled by subjects in the control group. They will differ slightly from difference of means (see text for details).

compatibility variable had to be replaced by a variable representing type of delayed cue.

The most important finding from the ANOVA of the unconditional recall data was a significant three-way interaction among immediate-test group, encoding task, and delayed-recall cue [$F(2,53) = 3.41$, $MS_e = .02$] that indicated that the two immediate-recall tests produced different patterns of enhancement. The same-level test produced facilitation in delayed recall relative to the control group only for the conditions in which the encoding tasks and the delayed cue were matched (i.e., the semantic encoding/semantic delayed-cue and the phonemic encoding/phonemic delayed-cue conditions; note that these were the compatible immediate-cue/delayed-cue conditions for the same-level test), but the different-level test produced facilitation in delayed recall primarily for the mismatched-encoding and delayed-cue conditions (note that these were the compatible immediate-cue/delayed-cue conditions for the different-level test). This effect boils down to the prediction from the cue-guided retrieval view. When immediate cues were compatible with delayed cues for the immediate-test groups (the compatible-immediate/delayed-cue cells in Table 2; first column of means), delayed recall was somewhat higher relative to the appropriate control cells (average improvement per cell was .04); however, when immediate cues were incompatible with delayed cues (Table 2, second column of means), delayed recall was lower, on average, relative to the appropriate control cells (average difference was −.02). Although the interaction pattern is consistent with the cue-guided retrieval view (outlined in the introduction), the magnitude of the

predicted compatibility effect appears to be slight. This issue is further explored below with an analysis of scores that measure the degree to which delayed recall was enhanced by immediate retrieval.

The ANOVA also showed that semantically encoded words were remembered better than phonemically encoded words [.25 vs. .12; $F(1,53) = 52.51$, $MS_e = .02$] and that delayed-recall cues that matched the original encoding level were more effective than delayed cues that mismatched the original encoding level [.22 vs. .15; $F(1,53) = 13.82$, $MS_e = .02$] (in Table 2, the matched-encoding/delayed-cue cells are represented in the "compatible" column for the same-level immediate-test group and in the "incompatible" column for the different-level immediate-test group; the mismatched-encoding/delayed-cue cells are represented in the alternative columns). Both of these effects parallel Fisher and Craik's (1977) immediate-recall findings.

The effects of immediate retrieval on memory could be somewhat masked by the above analysis because the overall proportion recalled does not distinguish between targets that were successfully retrieved on the immediate test and those that were not retrieved. To obtain another estimate of the influence of immediate retrieval across the different retrieval conditions, we used the "effectiveness score," *a*, employed in previous research on retrieval effects (see Bjork, Hofacker, & Burns, 1981; Lockhart, 1975). This score (developed by Lockhart, 1975, and extended by Bjork et al., 1981, to situations in which final testing involves recall) is designed to measure the degree to which delayed recall is enhanced by initial retrieval,

and is viewed as a relatively pure measure of enhancement that is free of potential problems encountered with other enhancement measures (Bjork et al., 1981).

The a score is based on a simple finite-state model that classifies items from a list into one of four states. The states represent whether or not the item is recallable in immediate recall (C or N) and whether or not it is recalled in delayed recall (C or N). An item in State CN is one that would be recalled in immediate recall (if immediate recall were permitted), but would not be recalled in delayed recall. Only items of this type can potentially be facilitated by the act of initial recall. Of current interest is the estimate of the probability, a, that items in State CN make a transition into State CC as a consequence of immediate recall being allowed to occur. The difference in final performance between a condition in which immediate recall is permitted (i.e., the immediate-test conditions; let P_T = final recall for the groups in which an immediate test is given) and a condition in which immediate recall is not permitted (i.e., the no immediate-test control group; let $P_{\bar{T}}$ = final recall of the control group) is thus a function of P_{CN}, the expected proportion of items in State CN (immediately recallable items not recalled on the delayed test) and a. Specifically, Lockhart (1975) derived this function as

$$P_T - P_{\bar{T}} = P_{CN}\left(\frac{a}{1 - a}\right).$$

Solving for a yields

$$a = \frac{P_T - P_{\bar{T}}}{P_{CN} + P_T - P_{\bar{T}}}.$$

Because the proportions on the right side of this equation are known (P_{CN} is observed in the conditions in which an immediate test is administered), a can be directly calculated. Note that even if the difference in final performance between the immediate-test condition and the control condition is near zero, a can still be substantially greater than zero as long as P_{CN} (the probability that an item will be immediately recalled but will not be recalled on the delayed test) is near zero (e.g., see Table 2 for robust a scores when the difference scores are near zero).

For the a-score analysis, we opted to define immediate-cue/delayed-cue compatibility as an independent variable so that the theoretically important effects due to compatibility would be manifested straightforwardly as a main effect. Note, however, that if the a-score analysis were computed with type of delayed cue as an independent variable (rather than with compatibility as an independent variable), the exact same pattern of effects would emerge. The pattern, however, would be more complicated to discern (from the tabled means) and describe because compatibility effects would be manifested as a three-way interaction, as was the case in the analysis of the unconditional recall data.

For each subject in each of the immediate-recall groups, we calculated a set of four a scores. These a scores represent performance in the four cells formed by the fac-

torial combination of encoding level and compatibility of immediate and delayed cues. In calculating these a scores, control-group recall was matched to the immediate-test groups' recall according to particular counterbalancing conditions. Thus, the difference between the immediate-test groups' final recall and the control group's final recall (this difference is used in the a-score calculation) for each particular encoding level and immediate-cue/delayed-cue compatibility cell was based on performance on the exact same set of stimulus items.[1] For 8 subjects in the different-level immediate-test group and for 1 subject in the same-level immediate-test group, two of the four a scores could not be calculated because these subjects failed to recall at least one item from a particular cell on the initial recall test (thus P_{CN} was not available). Because half of the data was not available for these subjects, they were discarded from the a-score analysis, leaving 14 subjects in the different-level group and 17 subjects in the same-level group. For 4 other subjects in the different-level group (3 in the phonemic encoding and compatible immediate-cue/delayed-cue cell and 1 in the phonemic encoding and incompatible immediate-cue/delayed-cue cell) and 6 other subjects in the same-level group (3 each in the phonemic encoding and compatible immediate-cue/delayed-cue, and phonemic encoding and incompatible immediate-cue/delayed-cue cells), there was only one missing score; consequently, these subjects were included in the ANOVA reported below by replacing the missing scores with the appropriate cell means (as shown in Table 2).

The a scores (see Table 2, last columns of the means) were analyzed with a three-factor mixed ANOVA, with type of immediate test as a between-subjects variable and encoding level and compatibility of immediate and delayed cues as within-subjects variables. There was a significant main effect of the cue-compatibility variable [$F(1,29) = 8.90$, $MS_e = .15$] such that compatible immediate and delayed cues produced significantly more enhancement in delayed recall than did incompatible immediate and delayed cues (.30 vs. .09). Note that this effect parallels the significant three-way interaction reported for the unconditional recall probabilities. No other significant effects were found. For what it is worth, a parallel analysis of the difference scores (between the immediate-test groups' final recall and the control-group's recall, i.e., the numerator of the a score) that included all of the subjects yielded the same pattern (refer to the means in parentheses in Table 2). Compatible immediate and delayed cues produced improvement, but incompatible immediate and delayed cues did not [.04 vs. $-.01$; $F(1,38) = 5.53$, $MS_e = .02$]. No other significant effects were found in this analysis.

The analyses of recall probabilities, a scores, and difference scores were generally consistent with the claim that the information incorporated into the memory representation of a target after initial recall, either in terms of the target's core encoding or in terms of retrieval structures (Whitten & Bjork, 1977), is information primarily rele-

vant to the initial retrieval cue. This conclusion is based on the finding that enhancement in delayed recall due to immediate retrieval was observed primarily when the delayed-retrieval cue cued the same attributes as those cued by the immediate cue (i.e., the compatible-cue conditions). The results from this first experiment can only be considered suggestive, however. Performance on the immediate tests for phonemically encoded words was so low that, for 19 immediate-test subjects, there was no immediate retrieval in at least one cell. Thus, the enhancing effects of immediate recall could not be ascertained in at least one cell for nearly half of the subjects in the immediate-test conditions. Because of this limitation, two additional experiments were performed to provide convergent data for the conclusion suggested by our initial results. As will be amplified below, the condition in which the immediate-recall cues cue the same type of information as that emphasized at encoding (same-level immediate test) is particularly revealing in ruling out at least one alternative to the idea that retrieval processing is closely guided by the retrieval cue. Accordingly, Experiment 2 was conducted to establish that the pattern of effects observed in the same-level immediate-test condition is reliable and is somewhat general across variations in the particular retrieval cues used.

EXPERIMENT 2

In this experiment, we focused exclusively on the mnemonic effects of immediate-recall tests that cued the same type of information as that emphasized at encoding (same-level immediate tests). We did so because the same-level immediate tests provide an opportunity to distinguish between the idea that retrieval processing is closely guided by the retrieval cue and the idea that retrieval serves to increase the variability of the initial encoding of a target regardless of the type of cue prompting retrieval (McDaniel & Masson, 1985). On the basis of the latter view, McDaniel and Masson (1985) predicted that delayed recall would be most enhanced when the delayed cues were *inconsistent* with the original encoding emphasis. In the present context (i.e., with a same-level immediate test), this circumstance arises only when the immediate cues are *incompatible* with the delayed cues because in these cases, the original encoding emphasis is inconsistent with the delayed cues. Thus, the variability notion of retrieval effects anticipates enhancement of delayed recall primarily in the incompatible immediate-cue/ delayed-cue cells, whereas the cue-guided notion of retrieval effects anticipates enhancement of delayed recall (through immediate retrieval) primarily in the compatible immediate-cue/delayed-cue cells.[2]

McDaniel and Masson (1985) reported evidence for the former pattern (i.e., the pattern consistent with a variability notion). To evaluate whether or not the discrepancy between this previous finding and the results of Experiment 1 (for the same-level immediate-test condition) might have been due to a difference in the particular items used

to cue recall, in the present experiment, the cues on the same-level immediate test and the delayed test were identical to those used by McDaniel and Masson.[3]

In addition to using a same-level immediate test (as in Experiment 1 but with exactly the same cues as those used by McDaniel & Masson, 1985), we used an immediate-recall test in which the cues were the same words that had been used for the encoding question (the identical-word test). For example, if the encoding question for the target word LARK was, "How well does the word rhyme with BARK?" the cue for LARK on the identical-word immediate test would be BARK. We included this identical-word immediate test to provide a broader context in which to examine the mnemonic effects of retrieval. We reasoned that, to the extent that retrieval is cue-guided, identical-word immediate cues (cues that are entirely redundant with the encoding context) would preclude the target encoding from becoming more variable due to immediate retrieval. That is, to the extent that a cue guides retrieval processing (keys the kind of information added to the trace due to retrieval), the identical-word immediate test should produce a clear-cut pattern of enhancement on the delayed test (as outlined above).

Alternatively, if the mnemonic effects of retrieval are somewhat independent of the immediate-retrieval cue (cf., Bartlett & Tulving, 1974; McDaniel & Masson, 1985), then results on the delayed-recall test reflective of a cue-guided retrieval process (enhancement in the compatible immediate-cue/delayed-cue conditions) would not necessarily obtain consistently, even for the identical-word/ immediate-test group. A strong version of this view is that retrieval, regardless of the nature of the retrieval cue, serves to recode a target's original encoding such that the variability of that encoding is increased (cf. McDaniel & Masson, 1985). If so, then even the identical-word immediate test (as well as the same-level immediate test) should produce a pattern of enhancement on the delayed test that is completely at odds with the pattern expected on the basis of a cue-guided retrieval process (as outlined above). A weaker version of this view is that retrieval will induce variability only when the cues are not identical to the encoding context. In such a formulation, it would be expected that the identical-word and same-level immediate tests would produce different patterns of enhancement.

Method

Subjects. The subjects were 48 male and female students enrolled in an introductory psychology course at the University of Notre Dame. Subjects received extra credit for their participation. Sixteen subjects were randomly assigned to each of the three experimental groups (identical-word immediate test, same-level immediate test, or no immediate test).

Materials and Procedure. The word list and acquisition task were identical to those in Experiment 1, except that some of the particular context words (used for the orienting task) were changed to conform with those used by McDaniel and Masson (1985). For immediate testing, two cued-recall tests were employed. One test (the same-level immediate test) was similar to that used in Experiment 1, in which the cues were associated with the targets at the same level (semantic or phonemic) at which the targets were encoded. The par-

ticular cues used were those used in McDaniel and Masson's Experiment 1. For the other test (the identical-word immediate test), the cues were the identical words used as the context words at encoding. For delayed testing, as in Experiment 1, half of the semantically encoded words were cued with a category member and half were cued with a rhyme, and the same was true for phonemically encoded targets. The delayed test was like the one used for delayed recall in Experiment 1, except that some of the particular cues were changed to conform with those used by McDaniel and Masson. For this test, the cues were words not used previously in the experiment (either in the encoding task or in the initial test task).

The procedure was the same as that in Experiment 1 except that, for the acquisition phase, the target words were presented for 7 sec each using a microcomputer. After completing the encoding task, depending on the group to which they were assigned, the subjects were either given one of the immediate cued-recall tests or were dismissed. Upon returning to the laboratory, 24 h later, the subjects received the delayed cued-recall test.

Results

Immediate recall. Table 1 shows the proportions of words recalled on the immediate-recall tests. A two-factor mixed ANOVA of these scores indicated that the identical-word test produced better performance than did the same-level test [$F(1,30) = 18.86$, $MS_e = .02$]. In addition, words that were semantically encoded and cued with a semantic cue were recalled better than words that were phonemically encoded and cued with a phonemic cue [$F(1,30) = 60.73$, $MS_e = .02$], and this effect tended to be more pronounced with the identical-word test than with the same-level test [$F(1,30) = 3.64$, $MS_e = .02$, $p < .07$]. This pattern of results parallels that reported by Fisher and Craik (1977, Experiment 3).

Delayed recall. The mean proportion of words recalled on the delayed test by the immediate-recall groups and

the no immediate-recall control group are displayed in Table 3 (first four columns of means). These scores were submitted to a three-factor mixed ANOVA, with type of immediate test (identical-word, same-level, or none) as the between-subjects factor and encoding task (semantic or phonemic) and compatibility of immediate and delayed cues as within-subjects factors (for the control group, the compatible cells were defined as those in which the delayed cue was matched with the level of the encoding task and the incompatible cells were those in which the delayed cue was mismatched with encoding-task level). Semantically encoded words were remembered better than phonemically encoded words [$F(1,45) = 37.19$, $MS_e = .02$] and the immediate-recall tests significantly increased delayed recall [$F(2,45) = 5.91$, $MS_e = .04$]. There was also a significant main effect of immediate-cue/delayed-cue compatibility [$F(1,45) = 22.30$, $MS_e = .03$], indicating that delayed-recall cues that were compatible with the immediate-recall cues (or with the encoding task for the control group) produced higher recall than did cues that were incompatible. The compatibility variable interacted significantly with immediate testing [$F(2,45) = 3.66$, $MS_e = .03$] such that the compatibility effect was much more evident for the immediate-test groups than for the control group.

To assess more directly the effects of immediate retrieval on delayed recall, we conducted an analysis of a scores as in Experiment 1. One subject in the identical-word immediate-test group and 1 subject in the same-level immediate-test group did not recall at least one item from one particular cell. These two missing values were replaced by the mean a score of the appropriate cell. The a score means are shown in Table 3 (last two columns

Table 3
Mean Proportion Recalled on the Delayed Test, Mean a Scores, and Mean Difference Scores in Experiment 2

| Encoding Task/ Immediate-Recall Cue | Same Level Immediate-Test Group | | | | Control Group | | | | a Scores and Difference Scores | | | |
	Compatible		Incompatible		Compatible		Incompatible		Compatible		Incompatible	
Semantic Encoding/ Semantic Cue	(s)	.44	(p)	.36	(s)	.26	(p)	.25	(s)	.29 (.18)	(p)	.20 (.10)
Phonemic Encoding/ Phonemic Cue	(p)	.30	(s)	.14	(p)	.19	(s)	.13	(p)	.32 (.11)	(s)	.04 (.01)
Mean		.37		.25		.22		.19		.30 (.14)		.12 (.06)
	Identical-Word Immediate-Test Group											
Semantic Encoding/ Semantic Cue	(s)	.46	(p)	.28	(s)	.26	(p)	.25	(s)	.37 (.21)	(p)	.04 (.03)
Phonemic Encoding/ Phonemic Cue	(p)	.35	(s)	.14	(p)	.19	(s)	.13	(p)	.47 (.17)	(s)	0.0 (0.0)
Mean		.40		.21		.22		.19		.42 (.19)		.02 (.02)

Note—Compatible = compatible immediate and delayed cues, Incompatible = incompatible immediate and delayed cues. (s) = semantic delayed cue, (p) = phonemic delayed cue. The control group received no immediate test. The set of control scores is reported successively to indicate the appropriate control value for each cell of the immediate-test groups. The a scores are based on Lockhart's (1975) measure (see text for full explanation). The difference scores are in parentheses and are based on the difference between proportion recalled by subjects given an immediate test and proportion recalled by subjects in the control group. They will differ slightly from difference of means (see text for details).

of means). A three-factor mixed ANOVA, with encoding level and compatibility of immediate and delayed cues as within-subjects factors and type of immediate test as a between-subjects factor, indicated that enhancement was significantly greater for the compatible immediate-cue/delayed-cue conditions than for the incompatible conditions [$F(1,30) = 34.87$, $MS_e = .08$]. This effect interacted significantly with type of initial test [$F(1,30) = 4.64$, $MS_e = .08$]. An examination of the a scores in Table 3 indicates that the compatibility effect was present in both immediate-test groups, and that for the same-level immediate-test group there was also noticeable enhancement in the incompatible cell when the immediate cue was semantic and the delayed cue was phonemic. There were no other significant effects in the a-score analysis.

An analysis of the differences (the means displayed in parentheses in Table 3) between unconditional recall in the immediate-test groups and the control group (i.e., the numerator of the a score) also showed that compatible immediate and delayed cues produced significantly better improvement than did the incompatible immediate and delayed cues [$F(1,30) = 16.88$, $MS_e = .03$]. The compatibility × initial test interaction was not significant ($F = 1.98$), although the mean differences did parallel the a scores.

Discussion

The results support the conclusion from Experiment 1 that the processing associated with initial retrieval is guided primarily by the type of retrieval cues present on the initial test. This idea is supported by the finding that the compatibility of the cues on the immediate test and the cues on the delayed test was an important determinant of the extent to which immediate retrieval enhanced delayed recall. As in Experiment 1, when the immediate cue and the delayed cue cued the same type of information (phonemic–phonemic or semantic–semantic), delayed recall was enhanced relative to the no immediate-test control group (average a score = .36). Delayed recall was not generally enhanced (average a score = .07), however, when the delayed cue was incompatible with the immediate cue. The implication, on the basis of the encoding-specificity principle, is that the type of information added due to initial retrieval (either in terms of retrieval routes or in terms of encoded features) is closely aligned with the nature of the initial cue.

Although delayed recall was generally enhanced only when the delayed cue was compatible with the immediate cue, enhancement was also observed for the same-level test subjects in the incompatible immediate-cue/delayed-cue cell in which the immediate cue was semantic. The reliability of this effect is examined in Experiment 3. One possible interpretation of this effect that is consistent with the view discussed in the preceding paragraph is that retrieval mediated by a semantic cue (on the immediate test) would induce semantic processing. Semantic processing should be generally beneficial for delayed retention (McDaniel & Masson, 1977), perhaps

even when nonsemantic cues are present in delayed testing under certain circumstances. This idea is pursued further in the following experiment.

EXPERIMENT 3

To reiterate, the primary finding is that delayed-recall cues that cue the same type of information as that cued on the immediate test (compatible immediate-cue/delayed-cue conditions) enhance recall relative to a no immediate-test control condition, whereas delayed-recall cues that cue information from a different level as that cued on the immediate test (incompatible immediate-cue/delayed-cue conditions) produce little or no enhancement in recall relative to the control condition. We have thus far favored an interpretation of this result that emphasizes two points. First, the kind of processing prompted by initial retrieval will be guided by the type of information (i.e., retrieval cue) used during retrieval. Second, the mnemonic benefit of initial retrieval will depend, in part, on the extent to which initial retrieval processing is reinstated in the delayed-recall task.

In the previous experiments, however, there may have been an additional factor that contributed to the obtained pattern of results. Consider the conditions for all of the immediate-test subjects in Experiment 2 and for some of the subjects in Experiment 1 (those in the same-level immediate-test group). An immediate cued-recall test was given in which all of the cues were compatible with the information that had been originally encoded about the target items. Then, 24 h later, another cued-recall test was given in which at least some of the cues (half) were completely compatible with the information originally encoded and the information provided on the immediate test. At this point, it would not be unreasonable for the subject to expect that the cues provided on the delayed test would be related to the initial conditions. Such an expectation might have produced a retrieval set that resulted in a misdirected search in the presence of the delayed cues that were incompatible with previous information. The idea here is that the poor enhancement associated with incompatible delayed-test cues may have been due partly to general retrieval strategies induced by the initial test experience (rather than to particular informational components strengthened by cue-guided initial retrieval). There is evidence from other paradigms that subjects can adopt particular retrieval strategies, and that doing so can significantly influence the kinds of memory performance that is obtained (Reder & Wible, 1984).

Another factor that clouds the interpretation of the outcome of the previous experiment is that the advantage of compatible immediate-cue/delayed-cue cells was systematically related to compatibility of the original encoding level and the delayed cue. If immediate-cue/delayed-cue compatibility effects were also to obtain after a different-level immediate test, however, then such compatibility effects (immediate-cue/delayed-cue compatibility) would not be confounded with encoding/delayed-cue compati-

bility. Although immediate-cue/delayed-cue compatibility effects were evident in the different-level immediate-test condition of Experiment 1, these results bear replication because of the number of subjects whose results had to be excluded from the enhancement-score analysis because of their failure to retrieve (on the immediate test) at least one item from several cells.

Accordingly, in the present experiment, we focused on both same-level and different-level immediate tests (i.e., the immediate-test conditions used in Experiment 1). In an attempt to offset any particular expectations about the relationship between the delayed cue and previous processing of the word, the subjects were forewarned prior to the delayed cued-recall test that the particular cues might or might not be related to the original way in which the subject thought about (i.e., encoded) the target words. With this instructional set, three kinds of results seemed possible.

If the results from Experiments 1 and 2 were due entirely to general retrieval strategies based on expectancies induced by the immediate-test experience and reinforced by some of the delayed-cue conditions, then the pattern obtained previously (summarized in the first sentence of this introduction) should not obtain. In particular, delayed cues that are incompatible with the immediate cues should not produce a decrement in enhancement effects relative to delayed cues that are compatible with the immediate cues. Alternatively, if the pattern of results that was obtained previously was not due to general expectancies induced by the immediate-test experience, then the present results should generally parallel the previous results.

There is a third, more complex possibility. It may be that general strategies for approaching the delayed test were induced in the previous experiments, and the pattern of delayed-recall results was therefore a combination of general expectancies (induced by the type of immediate test) operating along with cue-specific (immediate cues) retrieval effects. The idea here is that the retrieval processing guided by the immediate-retrieval cues had specific consequences for the memory system (either information related to the immediate cue or retrieval paths pertaining to the cued information was enhanced), but the particular pattern of enhancement in delayed recall was limited both by the particular cues used on the delayed test and by the expectations of how these cues might relate to previous encounters with the target. In this framework, in the current experiment, one might expect a somewhat different pattern of results from that obtained previously, but the pattern should still be related to cue-specific processing in immediate retrieval.

Specifically, on the basis of the reasoning developed earlier in this paper (hinging on the encoding-specificity principle), significant enhancement should still occur when the delayed cues are compatible with the immediate cues. However, if the delayed cue does not prompt recovery of the target, the subjects might engage in a search strategy less restricted by the particular cue (es-

pecially given the forewarning that the subjects received). Such a search might more nearly approximate free recall, which is more enhanced by semantic-level initial retrieval than by phonemic-level initial retrieval (Whitten, 1978). If this analysis is correct, we would expect that, regardless of immediate-test group (same-level or different-level), immediate retrieval would enhance delayed recall for *incompatible* delayed cues when the immediate retrieval is prompted by semantic but not by acoustic (phonemic) cues. To summarize, this last framework leads to the prediction that delayed recall will be significantly enhanced by immediate retrieval when delayed cues are compatible with immediate cues. In this paradigm, delayed recall will also be enhanced when immediate retrieval is prompted by semantic cues, even when the delayed cues are incompatible (i.e., when the delayed cues are rhymes). On the other hand, delayed recall will show little or no enhancement by immediate retrieval that is prompted by rhyme cues when the delayed cues are incompatible (i.e., when the delayed cues are at the semantic level).

Method

Subjects. The subjects were 48 male and female undergraduate students at the University of Notre Dame participating for extra credit in their introductory psychology course. Thirty-two subjects received an immediate-recall test, with 16 subjects assigned to the same-level test group and 16 assigned to the different-level test group. Sixteen subjects received no immediate test.

Materials and Procedure. The word list and the acquisition task were identical to those in Experiment 2. For immediate testing, two cued-recall tests were used. As in Experiments 1 and 2, the same-level test contained cues that were associated with the targets at the same level at which the targets were encoded. The different-level test, as in Experiment 1, contained cues that were associated with the targets at a different informational level than that at which the targets were encoded (i.e., semantically encoded targets were cued with a rhyme and phonemically encoded targets were cued with a category member). The particular cues used in both tests were those used in McDaniel and Masson's (1985) first experiment (thus, the same-level test was identical to the same-level test in Experiment 2). The delayed cued-recall test was the same type as those used in Experiments 1 and 2, with the particular cues identical to those used in delayed testing in Experiment 2.

The procedure was the same as that in Experiment 2, except for the following critical difference. As part of the instructions for the delayed cued-recall test, the subjects were informed that the cue word either could match the information emphasized in the original encoding of the target or could relate to the target in a manner not emphasized in the original encoding task. This idea was illustrated for the subjects with one of the encoding practice words (SONNET). The subjects were encouraged to use this general hint in trying to recall the target words.

Results

Immediate recall. A two-factor mixed ANOVA of the immediate-recall proportions, with encoding level as the within-subjects variable and type of immediate test as the between-subjects variable, showed that semantic encoding produced better recall then phonemic encoding $[F(1,30) = 58.03, MS_e = .01$; see Table 1 for means]. In addition, the same-level immediate cues produced higher recall than did the different-level cues, but only

for phonemically encoded items [means = .35 vs. .25 for phonemically encoded items and .51 vs. .54 for semantically encoded items; $F(1,30) = 4.99$, $MS_e = .01$].

Delayed recall. Delayed-recall performance was initially examined by analyzing the proportion of words recalled by the immediate-test groups and by the no immediate-test control group (Table 4 gives the means). A three-factor mixed ANOVA paralleling that performed in Experiment 1, with immediate-test experience as the between-subjects factor and encoding task and type of delayed cue as the within-subjects factors (refer to Experiment 1 for an explanation of why the analysis included type of delayed cues as a variable), showed that taking an immediate test improved delayed recall relative to not taking an immediate test [$F(2,45) = 5.43$, $MS_e = .05$]. An examination of these data (unconditional delayed-recall proportions in Table 4) indicates, however, that the pattern of improvement (on delayed recall) was nominally different after the same-level immediate test than after the different-level immediate test. The same-level immediate test produced improvement relative to the control condition in every cell except in the phonemic encoding/semantic delayed-cue cell (note that this is the *incompatible* immediate-cue/delayed-cue cell in which the immediate cue was phonemic); in this cell, improvement was only slight. The different-level immediate test produced noticeable improvement relative to the control condition in every cell except in the semantic encoding/semantic delayed-cue cell (note that this is the *incompatible* immediate-cue/delayed-cue cell in which the immediate cue was phonemic). This observation is supported by the marginally significant three-way interaction among encoding level, delayed cue, and immediate test [$F(2,45) = 2.80$, $MS_e =$

.02, $p < .08$]. Below, we report analyses of enhancement scores (*a* scores) and difference scores that further support the above observation with statistically significant effects.

Three other effects emerged from the present ANOVA. Semantically encoded words were recalled better than were phonemically encoded words [.33 vs. .18; $F(1,45) = 47.09$, $MS_e = .02$]; phonemic delayed cues produced higher recall in general than did semantic delayed cues [.30 vs. .21; $F(1,45) = 17.74$, $MS_e = .02$]; and delayed cues that matched the original encoding level were more effective than delayed cues that mismatched the original encoding level [.28 vs. .23; $F(1,45) = 6.43$, $MS_e = .02$].

As in the previous experiments, *a* scores were analyzed (see Table 4 for means). One subject in the different-level immediate-test group did not recall at least one item from two cells; consequently, this subject's scores were dropped from the analysis. Another subject in this immediate-test group and 3 subjects in the same-level immediate-test group did not recall at least one item from one particular cell; for these subjects, the missing values were replaced by the mean *a* score of the appropriate cell. A three-factor mixed ANOVA, with encoding level and compatibility of immediate and delayed cues as within-subjects factors and type of immediate test as a between-subjects factor, indicated that compatible immediate and delayed cues produced significantly more enhancement than did incompatible immediate and delayed cues [$F(1,29) = 23.97$, $MS_e = .11$]. No other significant effects emerged.

In the last view outlined in the introduction, the expectation was that the only cells in which enhancement would not occur would be the incompatible immediate-cue/delayed-cue cells in which a phonemic (rhyme) immediate

Table 4
Mean Proportion Recalled on the Delayed Test, Mean *a* Scores, and Mean Difference Scores in Experiment 3

Encoding Task/ Immediate-Recall Cue	Same Level Immediate-Test Group				Control Group				*a* Scores and Difference Scores			
	Compatible		Incompatible		Compatible		Incompatible		Compatible		Incompatible	
Semantic Encoding/ Semantic Cue	(s)	.41	(p)	.36	(s)	.24	(p)	.25	(s)	.44 (.17)	(p)	.18 (.11)
Phonemic Encoding/ Phonemic Cue	(p)	.27	(s)	.09	(p)	.18	(s)	.06	(p)	.45 (.08)	(s)	.03 (.02)
Mean		.34		.22		.21		.16		.44 (.12)		.10 (.06)
	Different-Level Immediate-Test Group											
Semantic Encoding/ Phonemic Cue	(p)	.44	(s)	.27	(p)	.25	(s)	.24	(p)	.46 (.20)	(s)	.09 (.05)
Phonemic Encoding/ Semantic Cue	(s)	.17	(p)	.32	(s)	.06	(p)	.18	(s)	.42 (.12)	(p)	.30 (.13)
Mean		.30		.30		.16		.21		.44 (.16)		.20 (.09)

Note—Compatible = compatible immediate and delayed cues, Incompatible = incompatible immediate and delayed cues. (s) = semantic delayed cue, (p) = phonemic delayed cue. For the *a*-score calculation, a reduced set of subjects in the immediate-test conditions was used (see text for details); the tabled mean proportion recall values reflect performance of all subjects. The control group received no immediate test. The set of control scores is reported successively to indicate the appropriate control value for each cell of the immediate-test groups. The *a* scores are based on Lockhart's (1975) measure (see text for full explanation). Difference scores are in parentheses and are based on the difference between proportion recalled by subjects given an immediate test and proportion recalled by subjects in the control group. They will differ slightly from difference of means (see text for details).

cue was used. Comparisons of each cell's enhancement score to zero (the value representing no enhancement) supported this expectation. Neither of the enhancement scores in the incompatible immediate-cue/delayed-cue cells in which a phonemic (rhyme) immediate cue was used were significantly greater than zero (a scores = .03 for same-level test and .09 for the different-level test; $Fs < 1$), and the enhancement scores for all of the other cells were significantly greater than zero [smallest $F(1,29) = 4.85$, $MS_e = .11$].

An ANOVA paralleling that reported above was performed on the difference scores between delayed recall after immediate testing and delayed recall after no-immediate testing (i.e., the numerator of the a score; see Experiment 1 for a detailed description of how these difference scores were computed). Note that for this analysis, no subjects have missing values. Corresponding to both the initial analysis of proportion recall and the a-score analysis, compatible immediate and delayed cues produced significantly better improvement than did incompatible immediate and delayed cues [$F(1,30) = 6.40$, $MS_e = .02$]. Moreover, as for the a scores, the difference scores were not significantly greater than zero in the incompatible immediate-cue/delayed-cue cells in which a phonemic immediate cue was used (largest $F = 1.47$), but the difference scores were significantly greater than zero in the incompatible immediate-cue/delayed-cue cells in which a semantic immediate cue was used (smallest $F = 5.04$).

Discussion

These results are important for several reasons. First, they demonstrate that even when subjects are forewarned about how the delayed-recall cues relate to the original encoding task, immediate-cue/delayed-cue compatibility is a potent determinant of the degree to which immediate retrieval enhances delayed recall. This is important because it implies that the compatibility effects are not simply due to a general expectation fostered by the immediate-test experience, an experience that, for example, might lead subjects who received the same-level immediate test to expect all the delayed cues to be similar to the original encoding context.

Second, immediate-cue/delayed-cue compatibility-enhancement effects were obtained for the different-level immediate-test group in a situation in which nearly all subjects recalled enough on the immediate test to provide an index of delayed enhancement (a scores). A similar demonstration in Experiment 1 may have been suspect because more than one third of the subjects in the different-level immediate-test group had to be excluded from the critical a-score analysis. This immediate-cue/delayed-cue compatibility effect for the different-level immediate-test group implies that the compatibility effect evidenced by the same-level immediate-test group was not due to congruence between the encoding level and the delayed cue (or to congruence between the encoding level and the immediate cue). If it were, then a different pattern of en-

hancement in delayed recall should have been obtained for the different-level immediate-test condition.

Third, there was some indication that the compatibility of immediate- and delayed-recall cues was not the only factor in determining the enhancement of an immediate retrieval on delayed recall, especially when the subjects were made aware of the possible relationships between the delayed-recall cue and the original encoding context. Another determining factor seems to be the type of immediate-retrieval cue, with semantic retrieval cues (but not phonemic retrieval cues) producing enhancement even when the delayed cue was incompatible with the immediate cue. This finding is consistent with those of previous studies showing that final free recall is most enhanced by initial recall cued by semantic cues (Bartlett, 1977; Whitten, 1978) and adds converging support to the idea that retrieval processing differs according to the type of cue prompting retrieval.

GENERAL DISCUSSION

Two ideas regarding the mnemonic effects of initial recall have been prominent in the literature. One idea is that the effects of initial recall are related to a common mechanism that operates similarly regardless of initial cue type (Bartlett & Tulving, 1974; McDaniel & Masson, 1985). The other idea is that the effects of initial recall are closely related to the nature of the initial recall cue (Bartlett, 1977; Craik, 1981; see also Runquist, 1983, p. 649, for similar ideas). The results of the present experiments support the latter hypothesis: initial cued recall appears to produce an elaboration of an existing memory representation, in terms of the core encoding itself and/or in terms of retrieval routes, that contains information related to the attributes emphasized by the initial cue. This is an important finding because it significantly extends the conditions under which mnemonic enhancement of initial recall has been found to be cue-related. Previous results had demonstrated differential enhancement (as a function of cue type) on final free recall when initial cued recall followed presentation of short lists (six or fewer items) within 15 sec and when final free recall was administered within 1 min after the completion of the set of short-list presentations (Bartlett, 1977; Whitten, 1978). The current findings show that cue-related mnemonic effects of initial recall can occur when initial cued recall follows relatively long lists (40 items) and when final cued recall is delayed for 24 h (using only compatible immediate-cue/delayed-cue conditions, Runquist, 1983, reported positive mnemonic effects of initial cued recall when final cued recall was delayed for up to 21 days). Another significant extension of previous findings was the use, in the present study, of final cued recall instead of free recall. Because cued recall is sensitive to the nature of the information encoded about the target (Fisher & Craik, 1977), the final cued-recall test provided a telling index for the kinds of attributes enhanced through initial recovery. Thus, the weight of the empirical and theoreti-

cal work (e.g., Bartlett, 1977; Bjork, 1975; Craik, 1981; Whitten, 1978) on retrieval processes is consistent with the idea that initial retrieval processing is guided by the retrieval cues provided, much as initial encoding processing is guided by a given orienting task.

What remains is to understand more fully the possible conditions under which initial retrieval processing does *not* appear to be closely tied to the provided retrieval cues. McDaniel and Masson (1985), using a paradigm almost identical to that used here, found that the immediate-retrieval cue had no effect on the delayed cued-recall pattern. One likely candidate for the discrepancy between our findings and those of McDaniel and Masson is that McDaniel and Masson manipulated the relationship of the immediate-recall cue to the original encoding level (semantic or phonemic) as a within-subjects variable. In the first and third experiments of the present study, the relationship of the immediate cue to the original encoding task was manipulated between subjects (that is, the same-level immediate test provided cues from the same level as that of original encoding and the different-level immediate test provided cues from a different level than that of original encoding). McDaniel and Masson's within-subjects manipulation of immediate-cue type (in terms of its relationship to original encoding level) may have created a retrieval context that masked or overrode other effects of the individual cues per se. One possibility is that the subjects' retrieval strategies (in the McDaniel & Masson study) were influenced by their exposure, on the immediate test, to cues that were consistent with the original encoding emphasis *and* to cues that were inconsistent with the original encoding emphasis. The third experiment of the present study suggests that when subjects are made aware (through instructions) that the recall cues can be either consistent or inconsistent with the original encoding, then the nature of the recall cue does not completely drive performance. Perhaps in this case, the subjects adopted retrieval strategies that were not as closely tied to the cues provided. If so, this could explain why McDaniel and Masson did not find evidence for cue-guided retrieval processing. More generally, subjects' retrieval strategies have been shown to significantly influence the kinds of effects obtained in other paradigms (Reder & Wible, 1984); thus, this may be a profitable direction for further exploration of the mnemonic effects of retrieval.

REFERENCES

BARTLETT, J. C. (1977). Effects of immediate testing on delayed retrieval: Search and recovery operations with four types of cue. *Journal of Experimental Psychology: Human Learning & Memory, 3*, 719-732.

BARTLETT, J. C., & TULVING, E. (1974). Effects of temporal and semantic encoding in immediate recall upon subsequent retrieval. *Journal of Verbal Learning & Verbal Behavior, 13*, 297-309.

BJORK, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 123-144). New York: Wiley.

BJORK, R. A., HOFACKER, C., & BURNS, M. J. (1981, November). *An "effectiveness-ratio" measure of tests as learning events*. Paper presented at the 22nd Annual Meeting of the Psychonomic Society, Philadelphia.

CRAIK, F. I. M. (1981). Encoding and retrieval effects in human memory: A partial review. In J. Long & A. Baddeley (Eds.), *Attention and performance* (Vol. 9, pp. 383-402). Hillsdale, NJ: Erlbaum.

DARLEY, C. F., & MURDOCK, B. B. (1971). Effects of prior free recall testing on final recall and recognition. *Journal of Experimental Psychology, 91*, 66-73.

FISHER, R. P., & CRAIK, F. I. M. (1977). Interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Human Learning & Memory, 3*, 701-711.

HANAWALT, N. G., & TARR, A. G. (1961). The effect of recall upon recognition. *Journal of Experimental Psychology, 62*, 361-367.

HOGAN, R. M., & KINTSCH, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning & Verbal Behavior, 10*, 562-567.

LOCKHART, R. S. (1975). The facilitation of recognition by recall. *Journal of Verbal Learning & Verbal Behavior, 14*, 253-258.

McDANIEL, M. A., FRIEDMAN, A., & BOURNE, L. E., JR. (1978). Remembering the levels of information in words. *Memory & Cognition, 6*, 156-164.

McDANIEL, M. A., & MASSON, M. E. (1977). Long-term retention: When incidental semantic processing fails. *Journal of Experimental Psychology: Human Learning & Memory, 3*, 270-281.

McDANIEL, M. A., & MASSON, M. E. J. (1985). Altering memory representation through retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 11*, 371-385.

REDER, L. M., & WIBLE, C. (1984). Strategy use in question-answering: Memory strength and task constraints on fan effects. *Memory & Cognition, 12*, 411-419.

RUNQUIST, W. N. (1982). Accessibility of information to extralist cues following periods of disuse. *Journal of Verbal Learning & Verbal Behavior, 21*, 563-577.

RUNQUIST, W. N. (1983). Some effects of remembering on forgetting. *Memory & Cognition, 11*, 641-650.

TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.

WENGER, S. K., THOMPSON, C. P., & BARTLING, C. A. (1980). Recall facilitates subsequent recognition. *Journal of Experimental Psychology: Human Learning & Memory, 6*, 135-144.

WHITTEN, W. B. (1978). Initial-retrieval "depth" and the negative recency effect. *Memory & Cognition, 6*, 590-598.

WHITTEN, W. B., & BJORK, R. A. (1977). Learning from tests: Effects of spacing: *Journal of Verbal Learning & Verbal Behavior, 16*, 465-478.

NOTES

1. Specifically, the difference was computed by subtracting the average control delayed recall for a particular counterbalanced set of items from the delayed recall of each immediate-test subject in that particular counterbalanced condition.

2. Note that with a different-level immediate test, the variability notion of retrieval effects anticipates enhancement of delayed recall primarily in the *compatible* immediate-cue/delayed-cue cells because it is in these cells that the delayed cues are inconsistent with the original level of encoding. Thus, enhancement effects produced by a different-level immediate test do not distinguish between the predictions derived from the variability notion and those derived from the cue-guided notion.

3. In Experiment 1, many of our cue words were different than those used by McDaniel and Masson (1985) because of our attempt to equate the baseline target-generation rate across cue types.