

Recency effects in direct and indirect memory tasks

WENDY A. MCKENZIE and MICHAEL S. HUMPHREYS
University of Queensland, St. Lucia, Queensland, Australia

In three experiments, subjects learned two lists under incidental conditions and were then given either a part-word or a word (extralist associate) cue. Each cue was related to one word in each list. Half the subjects were given production instructions (an indirect memory test), and half were given cued recall instructions (a direct memory test). When the interval between List 2 and the test was shortened, recency effects were found for part-word cues for both cued recall and production instructions. Little or no recency effects were found with word cues. These results are incompatible with a simple distinction between the types of memory trace or information that are tapped by direct as opposed to indirect memory tasks. Possible causes for the recency effect and for the difference between word and part-word cues are discussed.

According to Richardson-Klavehn and Bjork (1988), the tasks that Humphreys, Bain, and Pike (1989) refer to as cued recall and production can be classified as direct and indirect memory tasks, respectively. A direct memory task is defined by the instructions to subjects to utilize information about a particular episode. In the case of cued recall, subjects study a list of words and are then asked to retrieve those words with the aid of cues. The instructions specify the episode in which the target words have occurred and state the relationship between cues and targets. More specifically, in the case of part-word cues, the subjects are provided with cues that are partial representations of the target words (i.e., stems, endings, or fragments). They are told that each cue is part of a word from the target list, and they are asked to add letters to complete each cue with a word from the specified episode. If extralist word cues are provided (e.g., taxonomic category labels or preexisting associates of the target words), the subjects are told that each cue is related to a word from the target episode and should help them to recall the words from the list.

In contrast, an indirect memory task is defined by instructions that do not make any reference to a particular episode, but refer only to the task at hand. This classification includes the production task of interest in this paper. The term *production* is used by Humphreys et al. (1989) to refer to an indirect retrieval task. This type of task requires subjects, after studying a list of words, to produce

the first word that comes to mind in response to each cue; hence it is a production task.

Some of the differences between direct and indirect memory tasks have been so dramatic as to lead to proposals of separate memory systems (see, e.g., Tulving, 1985). For example, in experimental comparisons of the performance of normal and amnesic subjects, although amnesics have shown impairment on recognition, free recall, and cued recall tasks, they show normal facilitation of performance from a single study trial on stem completion and free association tasks (Graf & Schacter, 1985; Graf, Squire, & Mandler, 1984; Shimamura & Squire, 1984).

In addition to the differential effects of amnesia, Graf and Schacter (1987) have cited *levels of processing*, shifts in presentation modality and recency as variables that have been shown to differentially affect direct and indirect memory tasks. For example, Jacoby and Dallas (1981) showed that a levels-of-processing manipulation affected performance on a recognition task but not on a perceptual identification task. Kirsner, Milech, and Standen (1983) showed that a modality shift affected performance on perceptual identification but not on a recognition task. Similarly, Tulving, Schacter, and Stark (1982) found different retention functions for fragment completion and recognition. It would be wrong, however, to conclude from this kind of evidence that the instructional change is responsible for the effects. In these experiments, along with the vast majority of experiments in which direct and indirect memory tasks have been compared, type of instruction has been confounded with other task details such as type of cue (e.g., word vs. part-word) and type of output (e.g., the retrieval of a target word vs. a decision about a target word).

Indeed, some evidence suggests that with respect to modality shifts and levels of processing, type of cue is important. For example, when Nelson and McEvoy (1979) looked at modality shifts in cued recall, they found an effect of ending cues but not extralist word cues. Similarly,

The first two experiments constituted part of the first author's Master's thesis at the University of Queensland. They and the third experiment were supported by a grant from the Australian Research Council to the second author. We wish to thank J. Bain and D. Nelson for their helpful comments on the thesis, and P. Graf, M. Masson, L. Jacoby, B. Stein, and particularly J. Chumbley, for their helpful comments on an earlier draft of this paper. W. McKenzie is now at Monash University. Send reprint requests to M. Humphreys, Department of Psychology, University of Queensland, Queensland 4072, Australia.

Rundus (1977) showed that cued recall with an extralist taxonomic cue was far more sensitive to an intentional/incidental learning manipulation than was cued recall with a rhyme cue. M. S. Humphreys has confirmed that the latter observation is not specific to rhyming cues. In two experiments, the effects of semantic and nonsemantic orienting tasks on cued recall with an extralist taxonomic cue and on cued recall with an ending cue were compared. With word cues, the semantic orienting tasks produced higher levels of recall than did the nonsemantic orienting tasks. There was, however, almost no effect of the orienting task on recall when the target words were cued with an ending cue. Nelson, Canas, Bajo, and Keelean (1987) have also reported finding equivalent lexical set size effects with cued recall and production instructions. That is, the larger the number of English words that shared the same letters with the cue (word stems and fragments), the poorer the performance with both cued recall and production instructions.

Given these results regarding modality effects, levels of processing, and cue set size effects, it seemed appropriate to reexamine the effects of recency on cued recall and production. A paradigm used by Bilodeau (1967) seemed promising for this, because it permitted differentiation between recency effects produced by a retrieval process and recency effects produced by a postretrieval recognition check.

Bilodeau (1967) had subjects study two lists of words (L1 and L2) in succession. They were then given a cued recall test (with extralist associates) for L1 words. The two study lists, each of which had five words, were constructed from free association norms, so that for each cue, L1 contained the primary associate and L2 contained the secondary associate. On the average, under free association conditions, L1 responses to the cue were produced 32% of the time, and L2 responses to the cue were produced 24% of the time. The subjects were allowed 20 sec to study each list. The retention interval was filled with vowel-canceling activity and lasted either 2 or 7 min. The test instructions required the subjects to recall the L1 words, and a list of five cues to the left of five blanks was provided. The subjects' cues had the same order on the test list as that of the corresponding L1 and L2 targets on the study lists.

Somewhat surprisingly, performance increased over time. More L1 responses and fewer L2 responses were produced after 7 min than after 2 min. This increase in the number of L1 responses as the length of the L2-test interval increased resembles spontaneous recovery in proactive interference (PI) paradigms (Underwood, 1948). There is, however, a crucial difference. In the PI paradigms, subjects learn L1 and L2. After a retention interval, the subjects are tested on L2. In this design, an increase in the number of L1 responses could be due to a breakdown in list discrimination.¹ That is, with a short L2-test interval, the subjects might be relatively good at determining whether it was an L1 or an L2 response.

Thus, some L2 items might be retrieved but not produced, because the subject would recognize them as L2 items. If this ability to discriminate should get worse as the L2-test interval increases, responses that would have been suppressed with a short L2-test interval might no longer be suppressed. Thus, spontaneous recovery in a PI paradigm might result from changes in discrimination. In contrast, in the Bilodeau (1967) paradigm, if L2 responses were suppressed at the shortest L2-test interval, there would be an increase in the number of L2 responses as the L2-test interval increased. Thus, in the Bilodeau paradigm, the increase in the number of L1 responses and the decrease in the number of L2 intrusions as the L2-test interval increases does not result from changes in discriminability. Instead, they must reflect changes in the retrievability of either the L2 or the L1 response or both.

Thus, the Bilodeau (1967) paradigm seems to be well suited for investigating the comparability of the retrieval process in cued recall and production. However, predictions about performance in the Bilodeau paradigm depend on the extent to which a recognition check occurs and on just what produces recency. Furthermore, outside the context of a particular theory or framework, concepts such as retrieval are not well defined. Accordingly, the theoretical framework proposed by Humphreys et al. (1989) was adopted in order to specify more precisely the relevant constructs and to explicate possible patterns of results.

Unlike other theorists (e.g., Roediger & Blaxton, 1987; Tulving, 1985) of cued recall and production, Humphreys et al. (1989) drew a sharp distinction between cued recall with a list associate (the cue is studied with the target) and cued recall with an extralist associate (the cue elicits the target in free association or controlled association norms). The primary reason for the Humphreys et al. distinction is that they had noted that no matter how discriminative a contextual cue was, pairwise associations between cue and target and between context and target would not suffice for the learning of a crossed-associates paradigm. Instead, they proposed that an "interactive" association involving context, the list cue, and the list target was required. Humphreys, Wiles, and Bain (in press) have elaborated on this point by noting that information about the joint occurrence of the context, the cue, and the target is required to learn crossed-associates and AB-ABr paradigms. This three-way information is not required, and is probably not even available, when recall is cued with an extralist associate. Because of the similarities between cued recall with an extralist associate and production, Humphreys et al. (1989) then proposed that these two tasks involve the same pairwise cue-target association. To explain why there are far more extralist intrusions for production than for cued recall with an extralist associate, Humphreys et al. suggested two possibilities. The first was that for cued recall a retrieved response is subjected to a postretrieval recognition check. If this check fails, another retrieval attempt may be made.² The alternative was that subjects use the list cue and con-

text in an "additive" fashion. The latter point has been expanded on by Humphreys et al. (in press), who showed how pairwise associations between cues and a common target could be used to compute the intersection of the sets of targets produced by the two cues. In the case of cued recall with an extralist cue, this process would find the intersection between all the associates of the cue and all the items subsumed under the context cue. The computation of the intersection of the sets of targets produced by the two cues is a powerful retrieval mechanism, but it is still inadequate for an AB-ABr paradigm. Under both suggestions, cued recall with an extralist associate and production in response to the same cue involve the same cue-to-target association. In this sense, these two tasks can be said to involve a similar retrieval process.

In the Humphreys et al. (1989) theory, the explanation for cued recall and production with part-word cues differed in one respect from the explanation for word cues. That is, they proposed that an association between a modality-specific representation and a central representation of the same word was responsible for part-word cuing effects. In contrast, an association between two central representations was assumed to be responsible for cuing effects with extralist associates. The only implication that this has for the present series of experiments is the opportunity to strengthen the cue-target association on the study trial. If part-word cuing effects are mediated by an association between a peripheral and a central representation of the same word, there is an obvious opportunity to strengthen this association on the study trial. That is, it could be strengthened whenever the central representation is activated by the presentation of a physical stimulus. For example, it would be strengthened by having the subject read the word *cold* but not by generating the antonym to *hot* (see Jacoby, 1983). In contrast, the association between *needle* and *thread* may not be strengthened by the presentation of *thread* unless *needle* is generated or all possible associations are strengthened (see Humphreys et al., 1989; Humphreys & Galbraith, 1975; Humphreys et al., in press).

At this point, we can make some predictions about performance under cued recall and production instructions in the Bilodeau (1967) paradigm. The following predictions are shared by a theory that posits a postretrieval recognition check and the use of context to retrieve the intersection of the list and the cue's associative set and probably any other theory for direct and indirect memory tasks: (1) There should be more extralist intrusions with production instructions than with cued recall instructions; (2) with cued recall, as opposed to production instructions, there should be an increase in the number of L1+L2 responses; and (3) this increase in the number of L1+L2 responses may be due primarily to an increase in the number of L1 responses. The latter prediction depends on the extent to which the contextual retrieval cue and/or the recognition check permits discrimination between the L1 and L2 responses.

The recency predictions depend on how context is used and on the nature of the contextual cue. Before describing these predictions, however, we must define recency more precisely. Our measure of recency is based on the difference in the number of L2 and L1 words recalled (produced). The measure of recency is this difference for the condition with the short interval between L2 and test (the L2-test interval) minus the same difference for the condition with the long L2-test interval—that is, $(L2 - L1)_{\text{short}} - (L2 - L1)_{\text{long}}$. This difference between differences corresponds to the two-way interaction between test interval (short vs. long) and list (L1 vs. L2). The test for a difference between recall and production in this statistic corresponds to the three-way interaction of test condition (recall vs. production), L2-test interval (short vs. long), and list (L1 vs. L2). This measure of recency removes the expected main effect of more L1 than L2 responses (in Experiment 1, as in Bilodeau, 1967, the preexisting association was stronger for L1 than for L2). It also removes the expected main effect of more L1+L2 responses in recall than in production and the possible interaction with list (the advantage for recall may be greater for L1 than for L2).

In making predictions about recency, we first consider the possibility that only a postretrieval recognition check is used. Since the ability to discriminate should stay the same or get worse as the L2-test interval increases, a postretrieval recognition check may produce negative recency in this paradigm (e.g., if $L1_{\text{short}} = L1_{\text{long}}$ and $L2_{\text{short}} < L2_{\text{long}}$). Since Bilodeau (1967) found positive recency with cued recall instructions, it would appear that there is an effect on retrieval and that this effect is larger than the negative effect that might be produced by discrimination. Thus, we can predict that there will be positive recency effects in both cued recall and production, but that these effects might be somewhat smaller in cued recall. Note that under these assumptions, recency is produced by changes in the strengths of the cue to L1 and/or cue to L2 associations. One possibility is that the preexisting association is strengthened at study, and that this strengthened association either decays during the retention interval or is interfered with by the activity that occurs during this interval. With the additional assumption that the rate of loss is rapid at first and then slows down, the pattern of results found by Bilodeau (1967) can be explained.

These predictions about recency can change if context is used as a retrieval cue, as in the intersection model proposed by Humphreys et al. (in press). The critical question here is whether context changes during the L2-test interval. If not, the recency effect will still be produced by changes in the strength of the associations between the cue and the L1 or L2 response, and the previous predictions still hold. If, however, recency is produced by rapidly changing contextual stimuli (see Glenberg, Bradley, Kraus, & Renzalgia, 1983) or by changes in the strength of the context-to-target associations, a very different prediction

applies. That is, since context should at most play a minor role in production, there should be a substantially larger context effect in cued recall than in production.

EXPERIMENT 1

The procedures used by Bilodeau (1967) were adapted to include a production task, and part-word cues (stems and endings) were used in addition to extralist associate cues. Ending cues were used because they have been most prominent in cued recall research. Stem cues were also used in order to increase the generality of the findings and because they have featured more prominently in research on production. As in the study by Bilodeau, the study lists were constructed so that each of the test cues (stems, endings, or associates) was common to one word in each list. Unlike in Bilodeau, the order of the words in each study and test list was randomized in order to reduce the possibility that production subjects would notice that the words they were producing had come from the same list. Another change from Bilodeau was the use of incidental orienting tasks at study. The tasks—frequency and pleasantness ratings—had been used previously in an interference paradigm in which subjects learned an AB pair and then rated A for frequency or pleasantness (Humphreys & Bain, 1985). In this paradigm, subjects were less likely to recall B given A as a cue if they had rated A. It seems unlikely that interference would have occurred if subjects were recalling B while they were rating A. Thus, it was thought that the use of incidental orienting tasks should reduce the likelihood that subjects would retrieve and rehearse an L1 word while rating an L2 word.

Method

Subjects. Two hundred and forty students from an introductory course in psychology at the University of Queensland served as subjects for course credit.

Materials. The test booklets consisted of general experimental instructions, study and test instructions, two study lists, and one test list compiled according to the requirements of each experimental condition. When it was necessary, three puzzles were included as distractor activities.

Three sets of two study lists and one test list were constructed, one for each type of cue: stems, endings, or associates. Within each cue set, two study lists of 20 words each were constructed so that each cue was common to one word in each study list and L1 words had a higher probability of being produced in response to each cue than did L2 words. According to production norms (Graf, personal communication, 1986; Nelson, personal communication, 1987; Thomson, Meredith, & Browning, 1976), L1 words were produced an average of 31% of the time, whereas L2 words were produced between 14% and 15% of the time. Within each cue set, a test list consisted of 20 cues that were stems (the first three letters of a word), endings (the last three letters of a word), or associates. For example, if the cues were stems, the test cue was *sho*_, the L1 response was *show*, and the L2 response was *shoe*. If the cues were endings, the test cue was *_ench*, the L1 response was *wrench*, and the L2 response was *stench*. If the cues were associates, the test cue was *animal*, the L1 response was *dog*, and the L2 response was *cat*.

Design and Procedure. The subjects were tested in groups of up to 12. They were assigned to the cuing conditions (stems, end-

ings, or associates), to the L2-test interval conditions (short L2-test or long L2-test), and to the retrieval task conditions (cued recall or production) through random distribution of the test booklets.

All subjects received the same set of experimental instructions. They were told that the purpose of the experiment was to investigate the cognitive processes underlying various tasks. It was explained that they were to be presented with a number of different tasks, each on a separate page, and that not all subjects would be doing the same task at the same time. The subjects were told they would be given 3 min to complete each task, and they were asked to wait until they were requested to go on if they finished within the allowed time. All subjects were first given a puzzle to attempt in the initial 3-min period, in order to try to prevent them from guessing the true nature of the experiment.

After working on the puzzle for 3 min, the subjects studied the first list of words. A different L1 was used for each of the three types of cues. The subjects were asked to rate each of the L1 words in terms of how frequently they thought it occurred in the English language. The scale ranged from 1 (*very frequent*) to 5 (*rare*). They were instructed to indicate their judgment by placing their rating beside each word. Three minutes were allowed for completion of this task. Following this, the subjects were presented with the second list of words (again a different list was used for each of the three types of cues). The subjects were instructed to rate the L2 words according to how pleasant they thought each word was on a scale of 1 (*very pleasant*) to 5 (*very unpleasant*). They were allowed 3 min to complete the task.

Next, half the subjects received an immediate test, while the remaining subjects attempted another puzzle. That test was either a cued recall (for L1) or a production task. The subjects under cued recall instructions were asked to try to recall the words that they had made frequency judgments about—that is, the first list of words they saw. They were informed of the relationship between the cues that were provided and the L1 words. In the case of stems and endings, they were told to add the remaining letters to complete each cue in order to form any of the words about which they could remember having made a frequency judgment. In the case of associates, they were told to look at each cue to see if it reminded them of a word about which they could remember making a frequency judgment.

The subjects who received production instructions were given the same list of cues as were the subjects in the corresponding cued recall condition. In the case of stems and endings, the subjects were told to look at each cue and add the letters needed to make the cue into any meaningful English word, but to write the first word that came to mind. The subjects who received associate cues were told to write the first word that came to mind in the space provided, in response to each cue. Three minutes were allowed for the subjects to complete either their retrieval task or a puzzle.

The next 3-min session was filled with distractor activity, during which all subjects attempted a page of addition problems. Following this, the subjects who were in the immediate test condition attempted another puzzle. The subjects in the delayed test condition received either a cued recall (for L1) or a production task, as described above. Again, the subjects were allowed 3 min to complete their respective tasks. This design produced retention intervals between the presentation of L2 and test of either 0 or 6 min.

Results

Each subject received two scores: the number of L1 responses recalled and the number of L2 responses recalled (a maximum of 20 in each case). These results are shown in Table 1.

The number of L1 and L2 responses were subjected to a four-way analysis of variance. The between-group variables comprised type of test (cued recall or production),

L2-test interval (short L2-test or long L2-test), and type of cue (stem, ending, or extralist associate). List (L1 or L2) constituted the within-group variable. This analysis and an inspection of the results indicated that word and part-word cues produced different patterns of results. That is, there was a significant main effect of type of cue [$F(2,228) = 4.63, MS_e = 7.13$], and there was an interaction between type of cue and list [$F(2,228) = 6.30$]. In this and in all subsequent analyses, all results are significant at the .05 level, unless reported otherwise. Consequently, separate analyses were performed on the part-word and associate cues in order to obtain a clearer picture of the results.

Part-word cues. A preliminary analysis indicated that there was no significant effect of type of cue (stems or endings), nor was there any significant interaction with this variable. For this reason, type of cue was dropped from the subsequent analysis.

Significantly more L1+L2 responses were produced with cued recall (15.75) than with production (11.30). This main effect of type of test was reliable [$F(1,156) = 43.89, MS_e = 8.18$]. This increase in the number recalled (produced) was, however, greater for L1 responses than for L2 responses, as is shown by the significant interaction between type of test and list [$F(1,156) = 6.51, MS_e = 15.22$]. The cued recall subjects recalled 8.46 L1 responses and 7.29 L2 responses. The production subjects produced 5.13 L1 responses and 6.18 L2 responses.

The recency effect is shown in Table 2 for stems, endings, and associates. There is a positive recency effect for both cued recall and production with part-word cues. This result is supported by the significant interaction of L2-test interval and list [$F(1,156) = 9.06, MS_e = 15.22$].

The three-way interaction of type of test \times retention interval \times list response was not reliable [$F(1,156) = .36$]. This result indicates that the recency effect that occurs with part-word cues is similar for cued recall and production.

Associates. A separate analysis of the associate data resulted in a three-way analysis of variance. More L1+L2

Table 2
The Recency Effect [(L2-L1) short - (L2-L1) long] in Experiments 1, 2, and 3, as a Function of Type of Test and Type of Cue

Type of Cue	Cued Recall	Production
	Experiment 1	
Stems	3.25	2.40
Endings	3.05	1.80
Associates	-0.55	0.74
	Experiment 2	
Stems	2.95	6.55
Endings	4.05	4.00
Associates	2.75	0.00
	Experiment 3	
Stems	2.72	5.53
Associates	2.46	0.56

responses were produced with cued recall (14.13) than with production (10.33). Regardless of type of test, more L1 responses (7.39) were produced than L2 responses (4.84). These main effects of type of test [$F(1,76) = 28.14, MS_e = 5.13$] and list [$F(1,76) = 29.22, MS_e = 8.90$] were reliable. There was also a marginally significant interaction between type of test and L2-test interval. With recall, the number of L1+L2 responses was greater at the long interval (15.05) than at the short interval (13.20). With production, however, the number of L1+L2 responses was somewhat less at the long test interval (9.85) than at the short test interval (10.81). However, there was no evidence of a recency effect; neither the two-way interaction between L2-test interval and list nor the three-way interaction between these variables and type of test was significant [$F(1,76) = .01, MS_e = 8.9$, and $F(1,76) = .48, MS_e = 8.9$, respectively].

Extralist intrusions. Many more extralist intrusions were produced with production (8.31) than with cued recall (0.93). This main effect of type of test was significant [$F(1,228) = 386.84, MS_e = 8.44$]. This effect was present for both word and part-word cues, and it did not change from the immediate to the delayed tests. The average number of intralist intrusions (word cues only) was .59 and did not differ between cued recall and production.

Discussion

There were significant differences favoring the cued recall subjects in the total number of L1+L2 responses, and for part-word cues but not associates, this effect was greater for L1 responses. These predictions are common to both explanations proposed by Humphreys et al. (1989) for the differences between cued recall with an extralist word cue and production. In addition, with extralist word cues there was a tendency for recall of L1+L2 responses to be better at the short interval, whereas the reverse was true for production. This interaction is to be expected if a retrieved L2 response is occasionally suppressed because the subject recognizes it as being an L2 response. This should occur more often with the short L2-test interval, because list discrimination should be better. Thus, this

Table 1
The Mean Number of Responses from L1, L2, and Extralist Intrusions, as a Function of Type of Cue, Type of Test, and L2-Test Interval, in Experiment 1

Type of Cue		Cued Recall		Production	
		L2-Test Interval		L2-Test Interval	
		Short	Long	Short	Long
Stems	L1	7.30	9.10	4.80	4.55
	L2	7.70	6.25	7.50	4.85
	Extralist	.60	1.45	7.90	9.60
Endings	L1	7.75	9.70	5.15	6.00
	L2	8.15	7.05	6.65	5.70
	Extralist	1.25	1.80	8.10	7.70
Associates	L1	7.75	8.40	6.76	6.65
	L2	5.45	6.65	4.05	3.20
	Extralist	0.30	0.20	8.10	8.45
	Intralist	0.75	0.60	0.40	0.60

Note—The mean number of intralist intrusions for associates is also included.

interaction appears to indicate the operation of a post-retrieval recognition check. With respect to recency, there were clear effects with stems and endings but no evidence for an effect with associates. Although the recency effects with part-word cues were somewhat larger for cued recall than for production, this difference was not significant.

EXPERIMENT 2

In an effort to produce a more robust recency effect, a variation on Bilodeau's (1967) design was introduced. In Experiment 2, the L2-test interval was manipulated by holding the L1-test interval constant and manipulating the L1-L2 interval. That is, the short L2-test interval was produced by inserting an interval between L1 and L2 (L1-L2 test) and the long L2-test interval was produced by presenting L2 immediately after L1 (L1L2-test). In addition, the study lists differed from those used in Experiment 1. Instead of the L1 word's having a higher probability of being produced in response to each cue than the L2 word did, the strengths of the cue-target associations were equated across the two lists. In addition, which list was studied in the L1 position was counterbalanced across subjects.

These changes permit different hypothetical sources of recency to have an effect. In Experiment 1 and in Bilodeau (1967), rapidly changing contextual cues might have produced a recency effect, but interference from the distractor activity in the retention interval and decay might not have. The reason is that the distractor activity occurred after both L1 and L2 had been presented. In order for interference from the distractor activity or decay during the distractor interval to have had a larger effect on L2 than on L1, it must have been the case that recently strengthened associations were more susceptible to interference or decay. In Experiment 2, the distractor activity either occurred between the presentation of L1 and L2 or after the presentation of both. It was thought that if the distractor activity should produce more retroactive than proactive interference, or if there should be decay during the L1-L2 interval, there could be more recency in Experiment 2 than there had been in Experiment 1.

Method

Subjects. Two hundred and forty students from an introductory course in psychology at the University of Queensland served as subjects for course credit.

Materials. Three sets of two study lists and one test were used, one for each type of cue: stems, endings, and associates. The list-wide average strength for both L1 and L2 was about 22% for stems (Graf, personal communication, 1986), endings (Nelson, personal communication, 1987), and associates (Thomson et al., 1976).

Design and Procedure. The subjects were tested in groups of up to 12. They were assigned to the cuing conditions (stems, endings, or associates), to the retrieval task conditions (cued recall or production), and to the L2-test interval conditions (short L1-L2 test or long L1L2-test) through random distribution of the test booklets. All subjects received the experimental instructions that had been used in Experiment 1, and again the initial 3-min session was taken up with puzzle activity.

Next, subjects studied the first list of words. A different L1 was used for each of the three types of cues: stems, endings, and associates. The subjects were asked to rate the first list of words for frequency. This task was not paced, and the subjects were allowed 3 min to complete it.

The remainder of the tasks were presented in either of two sequences: (1) L1L2-test: subjects were required to rate L2 for pleasantness before completing three 3-min sessions of puzzle activity, followed by either a cued recall task for L1 or a production task. (2) L1-L2 test: subjects completed three 3-min sessions of puzzle activity before rating L2 for pleasantness, followed by either a cued recall task for L1 or a production task. This design produced an L2-test interval of either 0 or 9 min.

Results

The numbers of L1 and L2 responses are given in Table 3. These responses were subjected to a four-way analysis of variance. The between-group variables comprised type of cue (stems, endings, or associates), type of test (cued recall or production), and L2-test interval (short L1-L2 test or long L1L2-test). The within-group variable was the list (L1 or L2). Preliminary analyses had shown that the effect of the counterbalancing variable and all interactions with it were not significant.

This analysis and an inspection of the results indicated that word and part-word cues probably produced a different pattern of results. That is, the interaction of type of cue \times list approached significance [$F(2,228) = 2.84$, $MS_e = 14.95$, $p = .06$]. In order to obtain a clearer picture of the results, separate analyses of the L1 and L2 responses for part-word and for word cues were conducted.

Part-word cues. A preliminary analysis indicated that there was no significant effect of type of cue (stems or endings), nor was there any significant interaction with this variable, so it was dropped from the subsequent analysis. This resulted in a three-way analysis of variance.

Significantly more L1+L2 responses were produced with cued recall (15.10) than with production (11.59). This main effect of type of test was reliable [$F(1,156) = 29.79$, $MS_e = 7.59$]. This advantage tended to be greater

Table 3
The Mean Number of Responses from L1 and L2, and Extralist Intrusions, as a Function of Type of Cue, Type of Test, and Interlist Interval

Type of Cue		Cued Recall		Production	
		L2-Test Interval	Long	Short	Long
Stems	L1	5.75	7.55	3.45	5.20
	L2	8.60	7.45	9.95	5.15
	Extralist	1.15	1.15	6.60	9.40
Endings	L1	5.85	9.10	4.05	5.45
	L2	8.45	7.65	7.85	5.25
	Extralist	1.30	1.65	8.15	9.10
Associates	L1	6.80	8.20	5.50	5.20
	L2	8.25	6.90	5.90	5.60
	Extralist	0.40	0.35	7.60	7.65
	Intralist	0.30	0.85	0.70	0.75

Note—The mean number of intralist intrusions for associates is also included.

for L1 than for L2 as the interaction between type of test and list approached significance [$F(1,156) = 3.60$, $MS_e = 15.81$, $p = .06$]. The mean numbers of L1 responses were 7.06 and 4.54 for the cued recall and production subjects, respectively. The corresponding values for the L2 responses were 8.04 and 7.05, respectively. Second, regardless of type of test, more L2 responses (7.54) were produced than L1 responses (5.80). This main effect of list was significant [$F(1,156) = 16.74$, $MS_e = 15.81$]. It was qualified by two significant interactions. The interaction between type of test and the L2-test interval was significant [$F(1,156) = 9.89$, $MS_e = 7.59$]. With cued recall, the number of L1+L2 responses was greater at the long L2-test interval (15.53) than at the short interval (14.33). With production, the number of L1+L2 responses was less at the long interval (10.53) than at the short interval (12.65). In addition, the interaction of the L2-test interval and list [$F(1,156) = 25.48$, $MS_e = 15.81$] was also significant (see Table 2). This recency effect was similar for both cued recall and production with stem and ending cues, since the three-way interaction of L2-test interval \times type of test \times list [$F(1,156) = 1.23$, $MS_e = 15.81$] was not significant.

Associates. A separate analysis of the associate data resulted in a three-way analysis of variance. More L1+L2 responses were produced with cued recall (15.08) than with production (11.10). This main effect of type of test was significant [$F(1,76) = 27.14$, $MS_e = 5.82$]. Although the recency effect for cued recall was almost as large for associates as it had been for stems (see Table 2), there was no statistical evidence for a recency effect. Neither the two-way interaction between the L2-test interval and list [$F(1,76) = 1.43$, $MS_e = 13.21$] nor the three-way interaction of type of test \times L2-test interval \times list [$F(1,76) = 1.43$, $MS_e = 13.21$] was significant. A separate analysis of the cued recall results did not change this conclusion. The interaction between L2-test interval and list was not significant [$F(1,38) = 1.79$, $MS_e = 21.12$].

Intrusions. An analysis of the extralist intrusions confirmed previous findings that many more extralist intrusions were made with production (8.08) than with cued recall (1.00). This difference was significant [$F(1,228) = 383.77$, $MS_e = 7.84$]. This effect was present for both word and part-word cues, and it did not change as a function of the L2-test interval. The average number of intralist intrusions (word cues only) was .65, and it did not differ between cued recall and production.

Discussion

With cued recall, subjects produced more L1+L2 responses and far fewer extralist intrusions than they did with production. These effects were obtained with both part-word cues and extralist associates. In addition, with part-word cues there was a tendency for recall in the cued recall condition to be worse in the short than in the long L2-test condition, whereas the reverse was true for production. As previously noted, this small reduction in the number of L1+L2 responses in cued recall presum-

ably results from the operation of a postretrieval recognition check.

There was also a large positive recency effect with part-word cues, since the number of L2 responses relative to the number of L1 responses was greater with a short than with a long L2-test interval. Although the three-way interaction between type of test, L2-test interval, and list was not significant, the pattern was different from that in Experiment 1, in which the recency effect had been slightly larger for cued recall. In Experiment 2, the recency effect, at least for stems, seemed to be larger with production instructions.

Although the recency effect with associates was not significant, it should be noted that if there is an effect the pattern was very different than it was with part-word cues. That is, with associates, there was no sign of a recency effect with production instructions, whereas the magnitude of the effect with cued recall was almost as large as it was for stems.

EXPERIMENT 3

At this point we had found a recency effect with part-word cues but not associates. This failure to replicate Bilodeau (1967) was puzzling, so we looked for differences between our procedures and his. One difference was the nature of the activity during the L2-test interval. Bilodeau had used a vowel-canceling task, whereas our retention interval was filled with mazes, visual puzzles, and addition problems. It seemed possible that the verbal distractor activity used by Bilodeau had produced more interference with word associations than was being produced by our visual tasks. In Experiment 3, we replicated Experiment 2, using an auditory short-term memory task during the L2-test interval.

Method

Subjects. One hundred and sixty students from an introductory course in psychology at the University of Queensland served as subjects for course credit. They were randomly assigned to conditions, but the randomization procedure was not blocked. This resulted in unequal n s in the conditions. The n s ranged from 16 to 26. The subjects were tested individually.³

Design and Procedure. Two different cuing conditions (stems and associates) were used. The cues and lists were the same as those used for these conditions in Experiment 2. The instructions were the same as those used in the previous experiments, except for the changes required by computer presentation and recording. In addition, there was no initial puzzle prior to the presentation of L1. Each word in L1 was presented for 2 sec. Then the word disappeared, and a prompt appeared, asking the subject to make a response. After entering the rating, the subject pressed the return key and the next word appeared. After all 20 words in L1 had been presented, the subjects in the long L2-test condition (L1L2-test) were given instructions for the next rating task and proceeded to make their ratings. The subjects in the short L2-test condition (L1-L2test) were told that they were going to listen to a tape of words. Their task was to keep track of the last word, which began with an "s." The tape contained 180 words, which were presented at a 5-sec rate. After listening to the tape, the subjects wrote down the last word that began with an "s," and then they rated the L2

words. The subjects in the long L2-test condition listened to this tape after they had made their ratings on L2. The final task for both the short and the long L2-test conditions was a cued recall or a production test. This design produced an L2-test interval of either 0 or 6 min.

Results and Discussion

The numbers of L1 and L2 responses are given in Table 4. These responses were subjected to a four-way analysis of variance, as in the previous experiment. Preliminary analyses had shown that the effect of the counterbalancing variable and all interactions with it were not significant. There was a significant effect of type of cue [$F(1,152) = 4.13$, $MS_e = 8.00$] and a nearly significant interaction between type of cue, type of test, and list response [$F(1,152) = 3.73$, $MS_e = 12.60$, $p = .06$]. Accordingly, we proceeded to analyze stems and associates separately, as we had done in the previous experiments.

Stems. There was a significant effect of list [$F(1,75) = 6.40$, $MS_e = 14.59$], owing to more L2 (6.87) than L1 (5.33) responses. This main effect was qualified by two interactions. The interaction between type of test and list was significant [$F(1,75) = 8.15$, $MS_e = 14.55$]. For the cued recall subjects, the number of L1 responses (6.57) was somewhat greater than the number of L2 responses (6.37). For the production subjects, the number of L1 responses (4.09) was less than the number of L2 responses (7.38). In addition, there was a significant interaction between the L2-test interval and list [$F(1,75) = 11.43$, $MS_e = 14.55$]. This recency effect is shown in Table 2. Just as in the previous experiment, the recency effect is slightly larger for production. However, the three-way interaction between type of test, L2-test interval, and list was not significant [$F(1,75) = 1.31$, $MS_e = 14.55$].

Associates. There was a significant effect of test condition [$F(1,77) = 10.38$], owing to the greater number of L1+L2 responses for cued recall (12.77) than for production (10.85). There was also a significant effect of list [$F(1,77) = 9.79$, $MS_e = 11.79$], owing to the greater number of L2 responses (6.76) as opposed to L1 responses (5.05). There was also a significant effect of

the L2-test interval [$F(1,77) = 5.56$, $MS_e = 4.95$]. The number of L1+L2 words in the short interval (11.16) was less than the number in the long interval (12.46). There was, however, no interaction between the L2-test interval and list [$F(1,77) = 1.90$, $MS_e = 11.79$, $p = .17$], or between type of test, L2-test interval, and list [$F(1,77) = .75$, $MS_e = 11.79$]. Nevertheless, as can be seen in Table 2, the pattern of recency effects was essentially the same as in Experiment 2. That is, in both experiments, the recency effect for cued recall was almost as large for associates as it was for stems. With stems, the recency effect was even larger for production. However, with associates, the recency effect disappeared for production. Nevertheless, an analysis of the recall data failed to yield any statistical evidence for a recency effect. The interaction between L2-test interval and list was not significant [$F(1,40) = 1.80$, $MS_e = 16.71$].

Intrusions. An analysis of the extralist intrusions confirmed previous findings that many more extralist intrusions were made with production (7.96) than with cued recall (1.41). This difference was significant [$F(1,153) = 269.01$, $MS_e = 6.25$]. This effect was present for both word and part-word cues, and it did not change as a function of the L2-test interval. The average number of intralist intrusions (word cues only) was .68 and did not differ between cued recall and production.

GENERAL DISCUSSION

In all three experiments, and for both extralist associate and part-word cues, the number of L1+L2 responses was greater for cued recall than for production. This cued recall advantage also tended to be greater for L1 than L2. There was also a recency effect with part-word cues, which was at least as large with production as it was with cued recall. With associates, there was at best a marginal recency effect, which was, if anything, larger for cued recall than it was for production. In interpreting these results, we will first consider the question of whether the instructional manipulation worked. That is, were the production subjects behaving like recall subjects, or vice versa? Next we will consider the question of whether the retrieval process is similar for cued recall and production. Finally, we will interpret the results within the Humphreys et al. (1989) framework, in an effort to pinpoint the difference between extralist associates and part-word cues.

Did the Instructional Manipulation Work?

The primary procedure employed to prevent the production subjects from "catching on" to the relationship between the cues employed on the production task and the items in the lists was to embed these tasks in a long string of irrelevant tasks. This should have been most effective in Experiment 2, in which subjects experienced four separate filler tasks and two different orienting tasks before the production task. In addition, the fact that two study lists were used and that the order of the words on the study and test lists was randomized meant that subjects were

Table 4
The Number of Responses from L1 and L2, and Extralist Intrusions, as a Function of Type of Cue, Type of Test, and L2-Test Interval, in Experiment 3

Type of Cue	Response	Cued Recall L2-Test Interval		Production L2-Test Interval	
		Short	Long	Short	Long
Stems	L1	5.84	7.30	2.88	5.29
	L2	7.00	5.74	8.94	5.81
	Extralist	1.67	2.13	8.18	8.91
Associates	L1	4.69	7.31	4.00	4.19
	L2	6.69	6.85	6.94	6.57
	Extralist	.88	.96	7.28	7.48
	Intralist	.56	.67	.56	.91

Note—The mean number of intralist intrusions for associates is also included.

producing an essentially random mixture of L1, L2, and extralist intrusions. This should have made it difficult for them to notice that the responses they were producing had appeared together on a study list. In addition, the production subjects appeared to be behaving like production subjects who have not recently encountered associates. The normative probability of either an L1 or an L2 response is .45 for part-word cues and .46 for word cues. In the short L2-test conditions of Experiment 1 (production instructions), the probability of producing either response is .60 for part-word cues and .54 for word cues. These modest increases over the normative base rate do not suggest any substantial number of production subjects behaving like cued recall subjects.

The use of a normative baseline can provide only a rough guide to the amount of learning that occurred. However, an increase over an experimentally obtained baseline is a very robust effect with part-word cues, and production instructions and the typical amount of such an increase is broadly consistent with the increase over the normative baseline that we obtained. (Note also that the higher the baseline, the less room there is for an increase.) Furthermore, the significant interaction between L2-test interval and list that was found with part-word cues in Experiments 1, 2, and 3 clearly established that learning had occurred. It is only with word cues, where no such interaction was found, that we cannot be certain whether learning had occurred.

Furthermore, the cued recall subjects clearly behaved differently from the production subjects. Across all three experiments, with cued recall instructions, the probability of an extralist intrusion was .07 for part-word cues and .03 for extralist associate cues. The corresponding probabilities for production instructions were .42 for part-word cues and .39 for extralist associate cues. This shows that the cued recall subjects were clearly attempting to recall the words in the episode specified by the instructions. Their ability to do this, however, was limited. Across all three experiments, with part-word cues, the probability of either an L1 or an L2 response was .75 for cued recall and .57 for production. The corresponding probabilities with word cues were .70 for cued recall and .54 for production.

The cued recall subjects had even greater difficulty in limiting their recalls to the L1 response. This is revealed by the relative number of L1 responses [$L1/(L1+L2)$] produced or recalled in Experiment 2 (in this experiment, the normative cue-target strengths were equated for L1 and L2, and the list order was counterbalanced). These probabilities with part-word cues were .47 for cued recall and .39 for production. The corresponding probabilities with extralist word cues were .50 for cued recall and .48 for production. Note that if the production subjects had been attempting to recall, they would have attempted to recall the L2 responses, or both the L1 and L2 responses, but not just the L1 responses. Thus, the small difference between cued recall and production in the percentage of L1 responses does not result from production subjects'

behaving like recall subjects. Instead, it results from a failure of the recall subjects to limit their recalls to the L1 responses.

The recall subjects in Experiment 3 were better at limiting their responses to L1. With part-word cues, $L1/(L1+L2)$ was .51 for cued recall and .36 for production. The corresponding values for extralist word cues were .47 for cued recall and .38 for production. These changes from Experiment 2 to Experiment 3 may indicate that some subjects in Experiment 2 may have attempted to recall both L1 and L2, not just L1. Nevertheless, this possible failure to comply with the instructions did not change the pattern of recency effects.

Is the Retrieval Process Similar for Cued Recall and Production?

In Experiments 1, 2, and 3, with part-word cues, the proportion of L2 responses was greater for the short L2-test interval than for the long L2-test interval. This interaction (L2-test interval \times list) was significant in all three experiments, and the three-way interaction (type of test \times L2-test interval \times list) was never significant. Nevertheless, with stems in Experiments 2 and 3, recency was somewhat greater for production than for cued recall.

In all three experiments, with extralist associate cues, there was no statistical evidence for a recency effect. Neither the two-way interaction between L2-test interval and list, nor the three-way interaction between type of test, L2-test interval, and list was ever significant. Nevertheless, with recall, the magnitude of the recency effect in Experiments 2 and 3 was almost as large for extralist associates as it was for stems, and there was no sign of a recency effect with production. A conservative conclusion is that the magnitude of the recency effect is larger for part-word cues than for extralist associates, and if there is a recency effect with extralist associates, the pattern is almost certainly different than it is with part-word cues.

These results are clearly incompatible with a simple distinction between the type of memory trace or type of information tapped by direct as opposed to indirect memory tests (e.g., Roediger & Blaxton, 1987; Tulving, 1985). In the first instance, they require us to distinguish between the type of cue employed as well as the type of retrieval instruction. Second, they suggest that there are some important commonalities between direct and indirect memory tasks when type of cue is controlled. In particular, it looks as if the retrieval process may be similar. That is, when subjects are instructed to recall L1, a postretrieval recognition check should produce a negative recency effect in the short L2-test interval condition relative to the long L2-test interval condition. Furthermore, a postretrieval recognition check should never, or almost never, be employed with production instructions. Nevertheless, positive recency effects, presumably due to retrieval, were found for both cued recall and production with part-word cues.

A small negative recency effect due to a postretrieval recognition check is not, however, precluded. In all three

experiments, and for both types of cues, there was a tendency for the number of L1+L2 responses to increase from the short to the long L2-test interval for recall and to decrease for production. This interaction (type of test \times L2-test interval) was marginally significant for word cues in Experiment 1 and significant in Experiment 3. It was also significant for part-word cues in Experiment 2. This is the pattern that would be produced if subjects were better at discriminating L1 from L2 responses at the short L2-test interval and if this enhanced discriminability resulted in cued recall subjects suppressing some L2 responses they had retrieved. By itself, this process will reduce the number of L2 responses in the short L2-test condition without a corresponding rise in the number of L1 responses, and it will attenuate a positive retrieval-based recency effect. This could explain why the recency effect for stems was somewhat larger for production than for cued recall. It is also a reason for being cautious in interpreting the nonsignificant recency effects for cued recall with associative cues. That is, a very small negative effect due to a postretrieval recognition check could have prevented us from observing a significant positive effect due to retrieval.

A Theoretical Analysis

At this time, it appears that we can locate the recency effect with part-word cues in a retrieval process without making major theoretical assumptions. An explanation for the difference in recency with word and part-word cues, however, requires a more explicit theory or theoretical framework. As we have indicated, Humphreys et al.'s (1989) explanation for cued recall and production with part-word cues differed in one respect from their explanation for cued recall and production with extralist word cues. That is, with part-word cues the association was assumed to be between a modality-specific representation and a central representation, whereas with word cues the association was between two central representations. This distinction was originally introduced to explain why modality effects occur with part-word cues but not with word cues. Humphreys et al. (in press) have also noted that this distinction can be used to explain why cued recall with an ending cue is superior to cued recall with an extralist associate (see Nelson, 1989). The basic idea here is that when a subject reads or hears a word, there is an opportunity to strengthen the peripheral-to-central association. There is no obvious opportunity for a subject to strengthen an association between two central representations unless the subject "thinks" of one word when shown the other.

A similar explanation may apply to the recency effects found in these experiments. A recently strengthened peripheral-to-central association may be subject to decay or interference. In Experiments 2 and 3, this could occur during the L1-L2 interval. In Experiment 1, we would have to assume that the rate of loss was greater for the more recent list (L2) than for the earlier list (L1) or propose another explanation for the recency effect found in

that experiment. Because there would be a greatly reduced chance of strengthening the association between the central representations of the cue and the target when a target word is studied, the recency effect with extralist associates should be greatly reduced. Furthermore, the probability of strengthening the cue-target association when studying the target alone may be greater with the intentional learning instructions used by Bilodeau (1967) than with the incidental instructions used in these experiments. Note, however, that an additional explanation will be required if further research confirms that, with word cues, a larger recency effect is found for cued recall than for production. In the Humphreys et al. (in press) framework, this pattern would occur if the use of a contextual cue was responsible for the recency effect. There could be changing contextual cues as proposed by Glenberg et al. (1983), or a fixed contextual cue that was associated with every list item. Again, in Experiments 2 and 3 the activity in the L1-L2 interval could interfere with the associations between a fixed context and the L1 items.

As a final point, McKenzie (1990) used the design of Experiment 1 in two subsequent experiments with part-word cues. In both experiments, recency effects were substantially reduced and were in fact nonsignificant. However, when the results from those two experiments were analyzed along with the results from Experiment 1, significant recency effects were found for both stems and endings. We conclude that recency does occur with the Experiment 1 design, but that it is smaller than the effect found with the design used in Experiments 2 and 3. This suggests that the L1-L2 interval is an important component of the enhanced recency effects. Further research on the contents of this interval and the activity of the subject during it may illuminate the causes of recency still more.

REFERENCES

- ANDERSON, J. R., & BOWER, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*, 97-123.
- BILODEAU, E. A. (1967). Experimental interference with primary associates and their subsequent recovery with rest. *Journal of Experimental Psychology*, *73*, 328-332.
- GLENBERG, A. M., BRADLEY, M. M., KRAUS, T. A., & RENZALGIA, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*, 231-255.
- GRAF, P., & SCHACTER, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 510-518.
- GRAF, P., & SCHACTER, D. L. (1987). Selective effects of interference on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *13*, 45-53.
- GRAF, P., SQUIRE, L. R., & MANDLER, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 164-178.
- HUMPHREYS, M. S., & BAIN, J. D. (1985). Effect of retention interval on the cue repetition effect. *American Journal of Psychology*, *98*, 559-568.
- HUMPHREYS, M. S., BAIN, J. D., & PIKE, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review*, *96*, 208-233.

- HUMPHREYS, M. S., & GALBRAITH, R. C. (1975). Forward and backward associations in cued recall: Predictions from the encoding specificity principle. *Journal of Experimental Psychology: Human Learning & Memory*, *1*, 702-710.
- HUMPHREYS, M. S., WILES, J., & BAIN, J. D. (in press). Memory retrieval with two cues: Think of intersecting sets. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: A Silver Jubilee*. Cambridge, MA: MIT Press/Bradford Books.
- JACOBY, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, *22*, 485-508.
- JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *3*, 306-340.
- JACOBY, L. L., & HOLLINGSHEAD, A. (1990). Toward a generate-recognize model of performance on direct and indirect tests of memory. *Journal of Memory & Language*, *29*, 433-454.
- KIRSNER, K., MILECH, D., & STANDEN, P. (1983). Common and modality-specific processes in the mental lexicon. *Memory & Cognition*, *11*, 621-630.
- McKENZIE, W. A. (1990). *Recency effects in cued recall and production: Support for a common retrieval process*. Unpublished Master's thesis, University of Queensland, St. Lucia, Queensland, Australia.
- NELSON, D. L. (1989). Implicitly activated knowledge and memory. In C. Izawa (Ed.), *Current issues in cognitive psychology: The Tulane Floweree Symposium on Cognition* (pp. 369-388). New York: Erlbaum.
- NELSON, D. L., CANAS, J. J., BAJO, M. T., & KEELEAN, P. D. (1987). Comparing word fragment completion and cued recall with letter cues. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *13*, 542-552.
- NELSON, D. L., & McEVOY, C. L. (1979). Effects of retention interval and modality on sensory and semantic trace information. *Memory & Cognition*, *7*, 257-262.
- RICHARDSON-KLAVEHN, A., & BJORK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, *39*, 475-543.
- ROEDIGER, H. L., III, & BLAXTON, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & Cognition*, *15*, 379-388.
- RUNDUS, D. (1977). Maintenance rehearsal and single level processing. *Journal of Verbal Learning & Verbal Behavior*, *16*, 665-682.
- SHIMAMURA, A. P., & SQUIRE, L. R. (1984). Paired associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General*, *113*, 556-570.
- THOMSON, D. M., MEREDITH, S. G., & BROWNING, C. J. (1976). *Monash free association norms*. Monash University, Department of Psychology, Clayton, Victoria, Australia.
- TULVING, E. (1976). Ecphoric processes in recall and recognition. In J. Brown (Ed.), *Recall and recognition* (pp. 37-73). London: Wiley.
- TULVING, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385-398.
- 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, & Cognition*, *8*, 336-342.
- UNDERWOOD, B. J. (1948). "Spontaneous" recovery of verbal associates. *Journal of Experimental Psychology*, *38*, 429-439.

NOTES

1. At the short retention interval, the L2 memories might be stronger (i.e., more recognizable) than the L1 memories, and this difference in strength could be used as a basis for discrimination. With an increase in the L2-test interval, however, there should be, if anything, a smaller difference in strength, making discrimination more difficult and performance worse. Similarly, any theory of list discrimination based on list markers (Anderson & Bower, 1972) would also predict that list discrimination could only get worse as the L2-test interval increased.
2. This kind of postretrieval recognition check in which a consciously accessed response is evaluated is not the process critiqued by Tulving (1976). In that paper, Tulving was critiquing the idea that preconscious memories were evaluated. Jacoby and Hollingshead's (1990) proposal that postretrieval recognition checks are not always applied also serves to reconcile Tulving's (1976) arguments with the Humphreys et al. (1989) proposal.
3. In Experiment 3, we used individual testing rather than the group testing of Experiment 2. With individual testing, there seems to be less chance for confusion, because subjects are more conscientious and/or less inhibited about asking the experimenter for further details.

(Manuscript received September 19, 1989;
revision accepted for publication December 13, 1990.)