Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall

DOUGLAS L. HINTZMAN, DAVID A. CAULTON, and DANIEL J. LEVITIN
University of Oregon, Eugene, Oregon

Two experiments tested the hypothesis that the time course of retrieval from memory is different for familiarity and recall. The response-signal method was used to compare memory retrieval dynamics in yes–no recognition memory, as a measure of familiarity, with those of list discrimination, as a measure of contextual recall. Responses were always made with regard to membership in two previous study lists. In Experiment 1 an exclusion task requiring positive responses to words from one list and negative responses to new words and words from the nontarget list was used. In Experiment 2, recognition and list discrimination were separate tasks. Retrieval curves from both experiments were consistent, showing that the minimal retrieval time for recognition was about 100 msec faster than that for list discrimination. Repetition affected asymptotic performance but had no reliable effects on retrieval dynamics in either the recognition or the list-discrimination task.

Memory models commonly assume that a cue such as a word can trigger retrieval from memory of two kinds of information: a scalar *familiarity* value and structured *recall* (or recollection) of a previous event (Atkinson & Juola, 1973; Eich, 1982; Gillund & Shiffrin, 1984; Hintzman, 1986, 1988; Humphreys, Bain, & Pike, 1989; Murdock, 1982; Raaijmakers & Shiffrin, 1992). There is evidence from behavioral neuroscience that seems to support this distinction. Eichenbaum, Otto, and Cohen (1994) have proposed that the "parahippocampal" region of the medial temporal lobe mediates intermediate-term familiarity of individual stimuli, while the hippocampal formation and neocortex encode longer term associative information (see also Aggleton & Shaw, 1996).

Recent work from our laboratory using the response-signal method of Reed (1973, 1976) suggests that familiarity and recall processes may be characterized by different minimal retrieval times. The response-signal method varies the amount of time the experimental subject is given to make a binary decision about the memory probe and plots changes in memory performance as a function of processing time (e.g., Corbett & Wickelgren, 1978; Dosher, 1981, 1984a, 1984b; Dosher & Rosedale, 1991; Gronlund & Ratcliff, 1989; Hintzman, Caulton, & Curran, 1994; McElree & Dosher, 1993; Ratcliff & McKoon, 1989; Wickelgren, 1977, 1981; Wickelgren & Corbett, 1977). On each test trial, onset of the retrieval cue is followed by a variable lag, which ends with a signal that demands an immediate response regarding the test item's status. Subjects are trained to respond

at the signal, even if they have to guess. Typically, false-alarm rates and hit rates are about the same at very short lags, diverge suddenly as the lag increases, and approach asymptotic levels at lags of around 1,000–2,000 msec. An estimate of the point at which hits and false alarms diverge can be taken as the minimum time that it takes to process information relevant to the decision and to execute the response.

In our work, we have focused on within-experiment comparisons of these minimal times, or *intercepts*, across different manipulations and different retrieval tasks. Two previous studies are particularly relevant to the distinction between the hypothetical processes of familiarity and recall. In Hintzman and Curran's (1994) Experiment 3, subjects studied lists of nouns, some in singular and some in plural form (e.g., FROGS, TRUCK), and then were given response-signal tests requiring old–new judgments. The test items included studied target items, completely new words, and words that were targets with their pluralities changed (e.g., FROG, TRUCKS). In an inclusion condition, subjects were instructed to treat these plurality-changed words in the same way as targets, calling both "old"; and in an exclusion condition, they were instructed to treat them the same as new words, calling both "new." Hintzman and Curran found that the average intercept for discriminating targets from completely new words came 116 msec earlier in the retrieval episode than it did for discriminating on the basis of the target's plurality. Neither minimal time appeared to depend on whether the target word had been seen in the study list one or two times.

In Hintzman and Caulton (1997), after studying mixed lists of auditory and visual words, subjects were given response-signal tests in two retrieval tasks: yes–no recognition memory and judgments of presentation modality. For words that had a single study trial, minimal times for above-chance recognition averaged 119 msec (Experi-

ment 1) and 150 msec (Experiment 2) earlier than those for above-chance modality judgments. When words were presented three times in the same modality, the minimal time for modality judgments became shorter, but that for recognition memory showed no significant change. Recognition-memory intercepts also appeared to be unaffected by sensory matching at study and test—that is, although test stimuli were physically identical to visual study presentations, there was no evidence that visually studied words were retrieved more quickly than those that had been studied auditorily (see also Mulligan & Hirshman, 1995).

Taken together, these findings suggest three conclusions. First, information on whether a test item has occurred recently appears to be available earlier in the retrieval episode than is recall of details of the item's earlier presentation. Our tentative interpretation of this difference is that there are two underlying retrieval processes—an early one that we refer to as familiarity and a later one that we refer to as recall. Recalled or recollected information can, of course, contribute to old–new recognition judgments after it becomes available (this accounts for the difference between inclusion- and exclusion-condition retrieval functions), but we assume that the intercept for recognition memory is determined solely by the earlier, familiarity process. Second, the time at which a target word's familiarity first becomes available appears to be insensitive both to repetition and to the modality in which the word was studied. These (and other) variables may affect only the asymptotic level of familiarity. Third, however, dynamics of the recall of content (the target noun's plurality) and context (the target word's modality) may be affected differently by repetition during study. Such a difference would not be surprising, because recalling a word's plurality and modality must involve at least some processes that are different.

The evidence that retrieval dynamics of plurality and modality information are affected differently by repetition must, however, be considered weak. First, in the plurality-recall experiment, Hintzman and Curran (1994) compared effects of one versus two presentations, while in the modality-judgment study, Hintzman and Caulton (1997) used frequencies of one versus three. Second, the curve fits for the plurality-recall data of Hintzman and Curran (1994) were too unreliable to support confident acceptance of the null hypothesis, primarily because accuracy was generally low. Third, the exclusion task used by Hintzman and Curran (1994) was a kind of familiarity-recall hybrid in which subjects were instructed to reject as nontargets not only completely new words, which should be of low familiarity, but also words that seemed familiar but differed from targets in a small detail—the presence or absence of a terminal "s"—which presumably had to be recalled. The recognition and modality-judgment data of Hintzman and Caulton, by contrast, came from separate tasks. Subjects did not know at the time they studied a list which retrieval task would be required, but were instructed prior to some test lists to discriminate between old words and new words and prior to other test lists to discriminate between words that had been studied auditorily and visually. Old–new judgments could be based purely on familiarity, and modality judgments could be based purely on recall, but exclusion instructions seem to require a mixed decision strategy, and this difference in complexity could be responsible for the different results.

The purpose of the present study was not to answer such questions directly, but to further explore the retrieval dynamics of familiarity and recall—in this case, the recall of list context—using the response-signal task. In the experiments reported here, we compared retrieval curves for recognition memory with those for list discrimination (Anderson & Bower, 1974; Hintzman & Block, 1971; Hintzman & Waters, 1970; Winograd, 1968). In recognition memory, subjects must distinguish old words from new words, while in list discrimination they must distinguish whether a test word came from the immediately preceding list or the list before that. In Experiment 1, an exclusion task was used requiring acceptance of target-list words and rejection both of new words and of words from the nontarget list; this is analogous to the earlier experiment on recall of plurality (Hintzman & Curran, 1994). In Experiment 2, separate instructions were used to separate the recognition-memory and list-discrimination tasks; this is analogous to the earlier experiment on memory for modality (Hintzman & Caulton, 1997). Comparison of retrieval dynamics for familiarity and list discrimination across these two experiments might tell us something about the relationship between exclusion and source-judgment tasks.

EXPERIMENT 1

In this experiment, subjects studied a series of word lists, each of which was followed by a test list made up of response-signal trials. The words in a test list fell into the four cells defined by crossing presence versus absence in the immediately preceding study list (designated List 2) with presence or absence in the study list before that (List 1). Thus, each test list was made up of words that had occurrence once in List 1 and once in List 2, words that had occurred once in List 1, words that had occurred once in List 2, and words that had not occurred in either list. Words were sampled without replacement, so those in this last category were completely new to the experiment.

Just after each study list, the subjects were told whether the target list for the following series of test trials would be List 1 or List 2. When List 1 was the target, the subjects were supposed to reject new words and words that occurred only in List 2; when List 2 was the target, they were to reject new words and words that occurred only in List 1. It was assumed that words occurring in both lists could be discriminated from completely new words on the basis of recent occurrence or familiarity, but that discriminating list membership in the other two categories would require that contextual information be recalled. On the basis of this reasoning and the tentative conclusions from our earlier research explained in the introduction, we expected minimal times for discriminating on the basis of familiarity

to be shorter than those for discriminating on the basis of context or list.

Method

Subjects. Twelve University of Oregon undergraduates participated as subjects in six 1-h sessions for payment of $30.

Materials and Design. Stimulus materials were 2,544 nouns selected at random from the Hoosier Mental Lexicon (Nusbaum, Pisoni, & Davis, 1984). Words were assigned to conditions randomly and without replacement for each subject separately; they were also assigned randomly to within-list positions. An experimental session included a series of blocks, each consisting of a study list followed by a response-signal test list. A test list included four types of items, defined by crossing presence versus absence in the study list of the same block (List 2) with presence versus absence in the study list of the previous block (List 1). One word of each type was tested at each of six response-signal lags.

To anchor our terminology to the subject's perspective and to maintain consistency with the description of Experiment 2, we use the terms List 1 and List 2, with the understanding that the lists thus designated shifted from block to block—that is, what was List 2 in one block became List 1 in the following block. Thus, we refer to the four conditions as both-lists, List-1-only, List-2-only, and neither-list. Although all test lists had the same structure, there were two types of test condition, defined by the instruction given to the subject immediately prior to the test list. Both were exclusion instructions. The target List 1 instruction was to respond "yes" if the word had been in List 1 and to respond "no" if it was new or had been presented only in List 2. The target List 2 instruction was to respond "yes" if the word had been in List 2 and to respond "no" if it was new or had been only in List 1.

The structure of the study lists can be inferred from preceding description of the test lists. Each study list consisted of 24 words: 6 (the both-lists words for the current block) were shared with the previous study list, 6 (the both-lists words for the following block) were shared with the next study list, and 12 were unique to the current study list. Of this last group, 6 were List-2-only words for the current block and 6 were List-1-only words for the following block.

During the first session, the subjects were given 18 study–test blocks; during each subsequent session, they were given 19 blocks. The first block of Session 1 was used only for practice. The second block of Session 1 and the first block of each subsequent session served both as practice and to set up study presentations for the both-lists and List-1-only conditions of the following block. With results from these initial blocks excluded, data from 106 test lists were analyzed for each subject. For half of the test lists in each session, determined at random, the instructions designated the target list as List 1; for the other half, the target was List 2. Altogether, the design yielded 53 potential observations in each condition at each of the six lags, for each experimental subject.

Procedure. Stimulus presentation, millisecond timing, and response collection were controlled by an Apple Macintosh computer (Rensink, 1990). Study and test words were displayed on the computer screen in 24-point, lowercase Helvetica font. Instructions and feedback were displayed in 18-point, uppercase Helvetica font. Responses were made on the "Z" and "?/" keys of the computer keyboard, and hand assignments for the "yes" and "no" responses were counterbalanced across subjects.

Each study–test block proceeded as follows: First, the instruction, "Study list—press any key to begin" appeared in the center of the screen. After a keypress, the 24 study words were presented in the center of the screen at a 2-sec rate. At the end of the study list, the identity of the target list for the test series to follow—either "this list" (List 2) or "last list" (List 1)—appeared in the center of the screen with the instruction, "Press any key to begin test." A keypress was

followed 250 msec later by the first test word. After a lag of 100, 160, 250, 500, 1,000, or 2,000 msec, the test word disappeared, rows of asterisks appeared above and below the word's previous location, and an 800-Hz tone sounded for 32 msec.[1] The asterisks and tone were the signal to respond. After the response, the screen was cleared, the RT for that trial was shown in the middle of the screen for 500 msec, and the screen was cleared again. Following a 250-msec interstimulus interval, the next test word was displayed.

After all trials in a test list were completed in this way, feedback regarding speed and accuracy in that list was displayed for 2 sec. The speed measure was the percentage of trials with RTs under 300 msec, and the accuracy measure was the overall percentage of hits minus the percentage of false alarms. Subjects were given the double goal of attaining 90 on the speed measure and an accuracy score above 10. (Note that the accuracy score was based on all trials, including those with short lags, where accuracy is poor.) After this feedback, the next study–test block was initiated with the prompt to press any key to begin.

Results and Discussion

Three subjects failed to attend all six sessions, and their data were not analyzed. The data of 1 additional subject were dropped for apparent failure to follow instructions (performance in the List-1-target and List-2-target instruction conditions was the same). The following analyses are therefore based on 8 subjects. Trials in which the subject's reaction time (RT) after the signal was under 100 msec or over 350 msec (3% of the data) were excluded from all reported analyses.

Reaction times. In the response-signal procedure, RTs following the signal to respond typically drop over the shortest lags. This drop is assumed to reflect allocation of resources away from processing the retrieval cue and toward anticipation of the signal. An analysis of variance (ANOVA) on mean RTs collapsed over lags revealed no main effects but only a complex interaction, in which RTs to words that occurred only in the nontarget list were slower than RTs to words in the other conditions. This interaction, in turn, interacted with lag, in a way that is revealed by Figure 1. The top panel shows mean RTs for test lists for which instructions defined the target as List 1; the bottom panel shows RTs when the target was List 2. The biggest differences among conditions occurred at the 500-msec lag, where RTs for List-2-only items were unusually long when List 1 was the target and RTs for List-1-only items were long when List 2 was the target. These results suggest persisting difficulty in making a negative response to a familiar word, even after a 500-msec lag. This interpretation is consistent with the accuracy data, to which we turn next.

Hits and false alarms. To avoid hit rates of 1 and false-alarm rates of 0, the data of each subject were adjusted as recommended by Snodgrass and Corwin (1988). For each condition, the hit rate was defined as $H = (N_{hits} + 0.5)/(N_{targets} + 1)$ and the false-alarm rate as $FA = (N_{false\ alarms} + 0.5)/(N_{nontargets} + 1)$. Because response-signal RTs appear to include some processing time that can vary across conditions, performance is usually plotted as a function of the sum, lag + mean RT. This is done in Figure 2, which shows
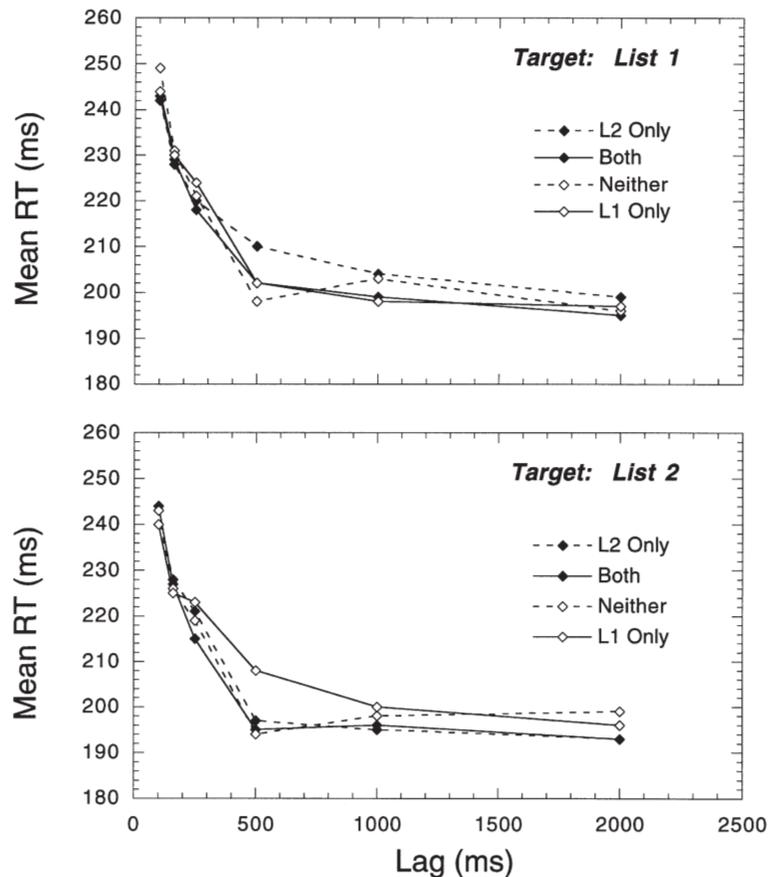
Figure 1. Mean reaction times following the signal to respond, Experiment 1. Top: List 1 target condition. Bottom: List 2 target condition. L1, List 1; L2, List 2.

H (rising curves) and *FA* (falling curves) separately for the two instructions and four word conditions. In accordance with the idea that familiarity is retrieved early in the retrieval episode, the ability to discriminate both-lists words from neither-list words was at chance only at the earliest lag and developed rapidly thereafter under both target-list instructions. Responses to the List-1-only and List-2-only words showed a more complex pattern. Under the List-1 instruction, across Lags 1–3, hits to List-1-only words were less frequent than false alarms to List-2-only words. The opposite result, however, did not occur when the target was List 2.

These results with test words occurring once in either List 1 or List 2 can be understood by assuming that List 2 words are more familiar because they were more recent, and that only familiarity is available during the earliest lags. If list discrimination requires information that is retrieved later, then instructions defining the target list should have no effect over these lags. To statistically evaluate this interpretation, a 2 (instructions) × 2 (lists) × 3 (lags) ANOVA was done on the List-1-only and List-2-only data from the shortest three lags. The proportion of "yes" responses was

significantly higher for List-2-only than for List-1-only [$F(1,7) = 7.45$, $MS_e = 0.014$, $p < .03$], and this difference interacted reliably with lag [$F(2,14) = 5.11$, $MS_e = 0.010$, $p < .03$]. No main effect or interaction involving the instruction manipulation approached significance. This statistical outcome is consistent with the view that information that becomes available early in the retrieval episode is not sufficient to discriminate between List 1 and List 2.

Discriminability. To plot discriminability against processing time in response-signal experiments, hit and false-alarm rates are usually transformed into a measure such as *d′*. In our work, we have used the logistic analogue of *d′*, referred to variously as d_L (Snodgrass & Corwin, 1988) and $2 \log(\alpha)$ (Macmillan & Creelman, 1991, 1996). This measure can be computed as

$$d_L = \ln\left[\frac{H(1-FA)}{(1-H)FA}\right]. \qquad (1)$$

Equation 1 can be used to obtain separate familiarity measures for each target-list instruction by contrasting hits in the both-lists condition with false alarms in the neither-list condition. If it is used for list discrimination, however
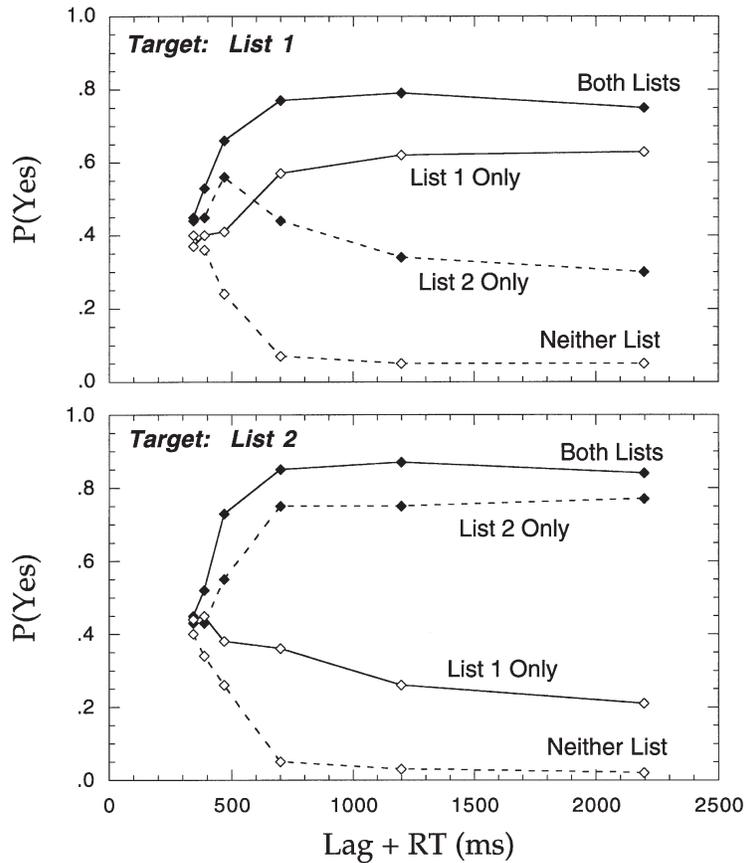
Figure 2. Probability of a "yes" response, Experiment 1. Top: List 1 target condition. Bottom: List 2 target condition. L1, List 1; L2, List 2.

[i.e., $H = P(\text{yes} \mid \text{target list only})$ and $FA = P(\text{yes} \mid \text{nontarget list only})$], the early inversion of H and FA seen in the top panel of Figure 1 will appear as negative d_L.

A more suitable measure that allows data from the two instruction conditions to be combined is based on the observation that Equation 1 is a log-odds measure (the natural logarithm of the ratio of correct responses to errors). As such, it can be generalized as follows:

$$d_L^* = \frac{1}{2}\ln\left[\frac{H_{T1}(1 - FA_{T1})H_{T2}(1 - FA_{T2})}{FA_{T1}(1 - H_{T1})FA_{T2}(1 - H_{T2})}\right]. \quad (2)$$

Here, the subscripts T1 and T2 refer to the instructions designating the target as List 1 and List 2, respectively. Equation 2, in effect, is the same as computing d_L for the two instruction conditions and averaging. (The factor, $\frac{1}{2}$, puts it on the same scale as Equation 1.) Thus, identifying H with both-lists and FA with neither-list yields a measure we associate with familiarity, and identifying H with target-list only and FA with nontarget-list only yields a measure of list discrimination. Applying Equation 2 in this way to the means of Figure 2 gives the two sets of data points that are shown in Figure 3.

Curve fitting. A consistent finding in response-signal studies has been that memory-retrieval data such as those plotted in Figure 3 can be closely described by a shifted exponential function—that is, an exponential approach to a limit that has its starting point or intercept displaced to the right:

$$\hat{d}_t = A\{1 - \exp[-R(t - I)]\} \text{ for } t > I, \text{ otherwise } 0. \quad (3)$$

Here, $\hat{d}$ is an estimate of the ability to discriminate targets from nontargets and t is time (lag + mean RT). The curve has three free parameters: I, the intercept or minimum time at which accuracy begins to rise above chance; A, the asymptote or terminal level of accuracy; and R, the rate at which the curve approaches asymptote. For data fitting, it is convenient to express the rate in terms of its inverse, $1/R$, which is measured in the same time units as I. Least squares fits of Equation 3 to the both versus neither and List 1 versus List 2 data are shown in Figure 3.

To assess differences among retrieval curves, Equation 3 was fitted separately to the both versus neither and the List 1 versus List 2 data of each subject, and the resulting parameter estimates were submitted to standard repeated measures statistical analyses. Summary statistics on individual
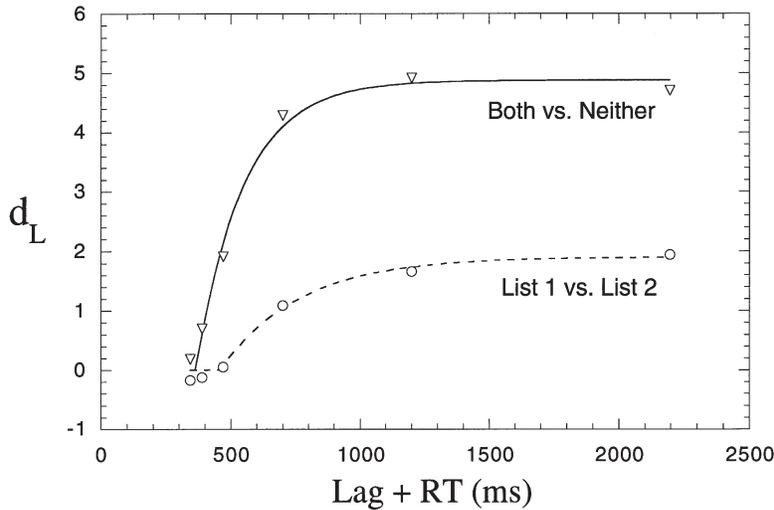
Figure 3. Discriminability curves from Experiment 1 for the both-list versus neither-list discrimination and for List 1 versus List 2.

subjects' best fitting curves are given in Table 1. For completeness, we show results of separate both versus neither fits for target-List-1 and target-List-2 instructions (d_L^* from Equation 1) and for both instructions combined (d_L from Equation 2). To avoid systematically negative d_L, just one fit is given for List 1 versus List 2, based on d_L^* (Equation 2). The right-hand column of the table shows mean goodness of fit in terms of the proportion of variance accounted for, adjusted for degrees of freedom (r^2; see Reed, 1976). The fits of Equation 3 to the individual discrimination functions were generally good.

Repeated measures t tests comparing the both versus neither curves for the two target-list instructions, as seen in the top two rows of Table 1, revealed no reliable instruction effects on any of the parameter estimates [$t(7) = 1.23$, $SE = 27$, for I; $t(7) = 1.36$, $SE = 36$, for $1/R$; and $t(7) = 2.22$, $SE = 0.35$, for A]. The difference in A is in a direction suggesting that subjects gave more weight to familiarity when the target was List 2 than when it was List 1; however, the difference only approached significance ($p = .062$). Comparison of the combined both versus neither curve with the List 1 versus List 2 curve, in the bottom two rows of the table, revealed significant differences in two of the parameters. The mean I was shorter for both versus neither than

for List 1 versus List 2 [$t(7) = 2.49$, $SE = .042$, $p < .05$], and the mean A was higher for both versus neither than for List 1 versus List 2 [$t(7) = 8.24$, $SE = 0.35$, $p < .001$]. The mean rate was faster ($1/R$ longer) for both versus neither than for List 1 versus List 2 , but this difference was not reliable [$t(7) = 1.62$, $SE = 65$, $p = .15$].

Overall, these data are consistent with the hypothesis that information sufficient to discriminate old words from new words is available earlier than the information needed for list discrimination. In keeping with our working hypothesis, the 115 ± 46 msec (mean $\pm$ SE) difference in intercepts is consistent with previous estimates of the difference in minimal retrieval times for familiarity and recall. Hintzman and Curran's (1994) data in comparing recognition memory with recall of a word's plurality yielded an estimated difference of 116 ± 30 msec.[2] Data for recognition memory versus modality judgments after a single study presentation have yielded estimates of 119 ± 31 msec (Hintzman & Caulton, 1997, Experiment 1) and 150 ± 49 msec (Hintzman & Caulton, 1997, Experiment 2). These values are all in the same range.

However, the exclusion task used in this experiment, like that of Hintzman and Curran (1994), was complex. Subjects were required to give "no" responses to neither-

Table 1
Summary Statistics for Fits of Shifted
Exponential Function to Data of Experiment 1

Curve	Measure	I (msec) M	I (msec) SE	$1/R$ (msec) M	$1/R$ (msec) SE	A M	A SE	Mean r^2
Both vs. Neither T1	d_L	388	22	122	28	4.42	0.25	0.93
Both vs. Neither T2	d_L	421	42	171	50	5.19	0.32	0.95
Both vs. Neither T1&T2	d_L^*	385	21	138	31	4.92	0.22	0.96
List 1 vs. List 2	d_L^*	500	45	244	69	2.03	1.29	0.87

Note—I, intercept; R, rate; A, asymptote; r^2, proportion of variance explained; T1, Target List 1 instruction; T2, Target List 2 instruction; d_L from Equation 1; d_L^* from Equation 2.

list items, which were completely new to the experiment, and also to nontarget-list-only items, which had acquired experimental familiarity through presentation in the non-target list. Moreover, the response demands for the latter items changed unpredictably from one test block to the next. Target List 1 instructions required "yes" responses to List-1-only words and "no" responses to List-2-only words, while target List 2 instructions reversed these assignments. It is possible that differences between the familiarity and list discrimination curves, and particularly the difference in intercepts, derives from this complexity rather than from a true difference in minimal retrieval times. Both versus neither decisions were easier than List 1 versus List 2 decisions at asymptote. If subjects handled the dual nature of the task by first assessing the test word's familiarity and then evaluating information on list membership only when familiarity was high, the dynamics parameters of response-signal curves could differ even if the two discriminations had the same minimal retrieval time.

Accordingly, in Experiment 2, we simplified demands on our subjects by dropping the exclusion instruction and using separate test lists for recognition judgments and for judgments of list membership. The separate-task procedure was thus similar to the one used by Hintzman and Caulton (1997) to compare retrieval dynamics of the recognition-memory and modality-judgment tasks.

EXPERIMENT 2

This experiment differed from Experiment 1 in several ways. First, subjects studied two lists in each block of the experiment, so each test list had its own List 1 and List 2. Second, following List 2, the subjects were instructed to prepare either for a series of recognition-memory tests or for a series of list-discrimination tests. Third, test lists for the two tasks had different structures. There were both old and new words in the recognition test lists, but only old words in the list-discrimination test lists. Fourth, some words were presented once and others were presented twice, but repetitions always occurred in the same study list. There were no words corresponding to the both-lists condition of Experiment 1.

One reason for using two presentation frequencies was to manipulate asymptotic accuracy in both the recognition-memory and list-discrimination tasks in order to better separate differences in retrieval dynamics from differences in asymptotic performance. As discussed earlier, previous studies found reliable repetition effects on the intercept for the modality-judgment task, but not for the recall of plurality and not for recognition memory (Hintzman & Caulton, 1997; Hintzman & Curran, 1994). In contrast to these null effects on the intercept, repetition increased asymptotic performance in all of the previously studied tasks. Effects of repetition on list-discrimination performance could be complex. Repeating a word is likely to establish its context in a particular list more firmly, but could also make the word seem more familiar and thus—to the

experimental subject—more likely to have been studied in List 2 (Hintzman & Waters, 1970; Winograd, 1968).

Method

Subjects. Thirty-three subjects were recruited from the University of Oregon community to serve in five experimental sessions. Of these subjects, 6 failed to complete all the sessions, and their data were dropped, leaving a total of 27. The subjects were paid $5 per session, but received an extra $5 if any session lasted more than 1 h.

Materials and Design. Stimulus words were taken from the same pool as in Experiment 1. Each block of the experiment consisted of a study List 1 and a study List 2, separated by a mathematical estimation task, and a response-signal test list that was administered immediately following List 2. In half of the blocks, the test list required old–new recognition decisions; in the other half, it required discrimination of list membership.

Each study list included 7 words that appeared once and 7 that appeared twice, for a list length of 21, all randomly ordered. Each list-discrimination test list consisted of the 28 old words from the preceding List 1 and List 2. Two words fell in each cell of a 2 (lists) $\times$ 2 (frequencies) $\times$ 7 (lags) design. Each recognition test list was exactly like a list-discrimination test list, but with the addition of 14 new words, 2 of which were tested at each of the seven lags. A recognition test list thus consisted of 42 trials. Words were randomly sampled and assigned to conditions without replacement. In both types of test list, conditions were ordered randomly.

The math estimation task that filled the interval between List 1 and List 2 was designed to prevent rehearsal but not to otherwise interfere with memory for the words. The math problems required estimating—not calculating—products and squares of random integers between 10 and 30. To enforce the restriction against explicit calculation, the problems were presented at a rate of one every 2 sec.

Procedure. The apparatus was the same as in Experiment 1. The two response alternatives were "old" versus "new" in the recognition-memory task and "List 1" versus "List 2" in the list-discrimination task. The "Z" and "?/" keys were used for both tasks, with assignments counterbalanced orthogonally across subjects to ensure that the key for "old" corresponded to the key for "List 1" as often as to that for "List 2." On the first day, to practice responding within 300 msec of the signal and to learn the response mappings, subjects were given one training block on each task. There were 8 data-gathering blocks on the first day and 12 blocks on each subsequent day, none of which were for practice.

Each study–test block proceeded as follows: First, "Press any key to begin first list" appeared in the center of the computer screen. Following a keypress, the study words of List 1 were presented for 2,000 msec each. After the 21st word, the instruction, "Please estimate the following values" appeared on the screen, and this was followed by 12 math estimation problems for 2,000 msec each. Subjects made their estimates using the numeric keypad of the computer. After the 12th problem, "Press any key to begin second list" appeared, and List 2 was presented just as List 1 had been. The 21st word of List 2 was replaced with a screen announcing the type of test list that would follow (either "old/new" or "List 1/List 2"). When a key was pressed, another screen instructed the subject to "Press any key to begin test"; an additional keypress initiated the sequence of test trials. The response-signal testing procedure was the same as in Experiment 1, except that the lags were 100, 140, 200, 300, 500, 1,000, and 2,000 msec.

Results and Discussion

Our primary interest was in curve fits for the four conditions defined by crossing recognition versus list discrimination with frequencies of 1 versus 2. Fits of Equation 3 to response-signal data are sometimes so poor as to
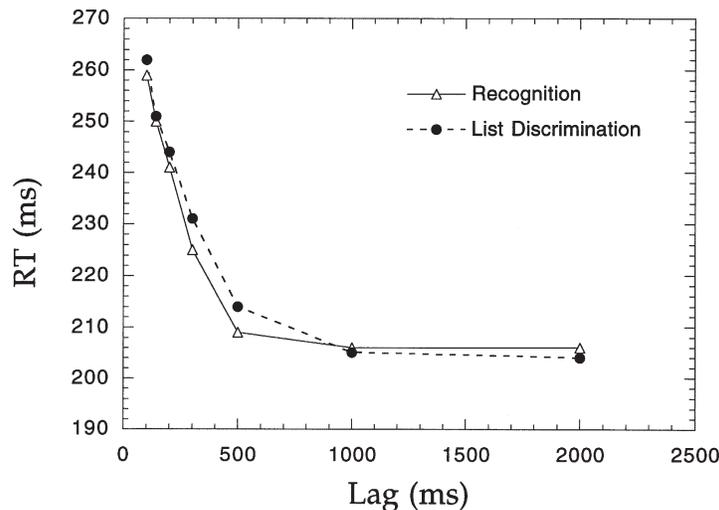
Figure 4. Mean RTs following the signal to respond, Experiment 2.

yield parameter estimates that are unreliable, and (because Equation 3 is nonlinear) possibly biased. For this reason, in our laboratory we routinely discard data of any subject with a curve fit that does not account for at least 50% of the variance ($r^2 < .50$). Nine of the 27 subjects who completed all sessions had at least one curve fit that failed to meet this criterion, and their data were dropped. Data were also dropped for an additional subject with acceptable r^2 values, but with one intercept estimate that was an outlier ($z = 3.39, p < .01$, according to Table VIIIa of Barnett & Lewis, 1984). To allow repeated measures statistics to be performed, all the data from these 10 subjects were dropped. For interested readers, statistics on the deleted subjects' curve fits are presented in the Appendix. All data and analyses that follow are from the remaining 17 subjects. As in Experiment 1, trials with RTs under 100 msec or over 350 msec (about 5% of the data) were excluded.

Reaction times. An analysis of recognition RTs collapsed over lag showed that responses to old items averaged 3.2 msec faster than responses to new items [$t(16) = 4.01, SE = 0.8, p < .01$]. This old–new difference may be a consequence of base rates, since two thirds of the recognition test words were old. A second analysis was done comparing RTs to old items only. Factors of the ANOVA were 2 tasks (recognition vs. list discrimination) $\times$ 2 lists (List 1 vs. List 2) $\times$ 2 frequencies (1 vs. 2 study trials). RTs averaged 2.9 msec faster in the recognition-memory task than in the list-discrimination task [$F(1,16) = 5.81, MS_e = 48.6, p < .05$], 1.9 msec faster to List 2 than to List 1 [$F(1,16) = 23.66, MS_e = 5.4, p < .001$], and 1.5 msec faster to frequency = 2 words than to frequency = 1 words [$F(1,16) = 10.35, MS_e = 7.1, p < .01$]. There were no reliable interactions.[3]

A simple interpretation of these small but reliable effects is that differences in difficulty are reflected in response latency. As in Experiment 1, the RT differences were most pronounced at intermediate lags, where difficulty is most

likely to be a factor. Figure 4 shows how the largest of these effects—the difference between tasks—interacted with lag.

Hits and false alarms. As in Experiment 1, H and FA for each subject were computed so as to avoid values of 0 and 1 (Snodgrass & Corwin, 1988). Figure 5 shows mean H and FA values in the recognition-memory task. At the shortest lag, all such values were in the neighborhood of 0.7, reflecting subjects' sensitivity to base rates (two thirds of the test words were old). As lag + RT increased, FA rates to new items dropped and H rates to old items approached levels reflecting both their frequency and their recency (i.e., frequency = 1 < frequency = 2, and List 1 < List 2).

Figure 6 shows H and FA for the list-discrimination task. Here, P(List 2), the probability of a "List 2" response, was arbitrarily selected as the ordinate. P(List 1) is its complement [i.e., P(List 1) = 1 − P(List 2)]. Ideally, repetition would shift the List 2 curve upward and the List 1 curve downward, but these data show little or no effect of repetition on list identification of words from List 1. This pattern can be explained by assuming that repetition had two effects: Repeating words improved recall of their context or list membership, but also increased their familiarity, and subjects treated high familiarity as evidence that a word had come from the more recent list. In the case of List 2 words, these two effects of repetition would be mutually reinforcing, but in the case of List 1 words, they would tend to cancel each other out.

We have no strong evidence to support this interpretation, but it is generally consistent with our other results. For example, Figure 5 suggests that recency and frequency are both associated with high familiarity. Also, although the difference was not statistically significant, comparison of the both versus neither asymptotes of Experiment 1 suggested that subjects relied more on overall familiarity when the target list was List 2 than when it was List 1.

Discriminability and curve fitting. For both tasks, Equation 1 was used to transform values of H and FA into
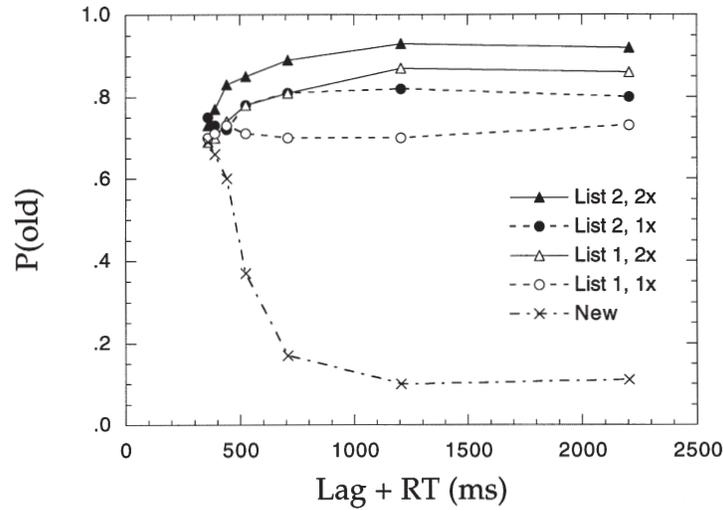
Figure 5. Probability of identifying a test item as "old" in recognition testing, Experiment 2. 1×, one presentation; 2×, two presentations.

d_L. Figure 7 shows separate discrimination data for each task and each frequency, derived from the average H and FA of Figure 6. Each data set has also been fitted with the shifted exponential function (Equation 3). The average fits suggest that repetition increased asymptotes without affecting the intercepts. However, both intercepts and rates separate according to task.

The four retrieval curves were fitted separately for each subject, and 10 subjects were deleted according to criteria discussed earlier (see the Appendix). Table 2 shows summary statistics of the best fitting parameters for the 17 subjects whose data were retained. The individual-subject estimates of the I, $1/R$, and A parameters were submitted to three separate repeated measures ANOVAs, each with two levels of task (recognition vs. list discrimination) and two levels of frequency (1 vs. 2). Recognition-memory intercepts were reliably shorter in the analysis of I than in list discrimination [$F(1,16) = 14.49$, $MS_e = 11,433$, $p < .005$]. Neither the effect of frequency nor its interaction with task approached significance. Subsequent t tests confirmed that the task difference was significant at both levels of frequency. In the analysis of $1/R$, rates tended to be slower ($1/R$ longer) at frequency = 2 than at frequency = 1 [$F(1,16) = 2.77$, $MS_e = 17,181$, $p = .12$], but neither this nor the other effects were statistically reliable. In the analysis of A, there were main effects of both task and frequency [both $Fs(1,16) > 18$, $p < .001$], but there was no reliable interaction.
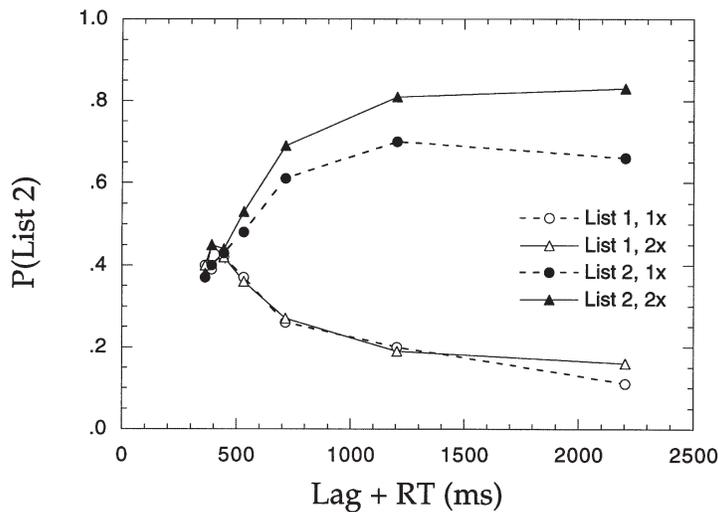

Figure 6. Probability of identifying a test item as belonging to List 2 in list-discrimination testing, Experiment 2. 1×, one presentation; 2×, two presentations.
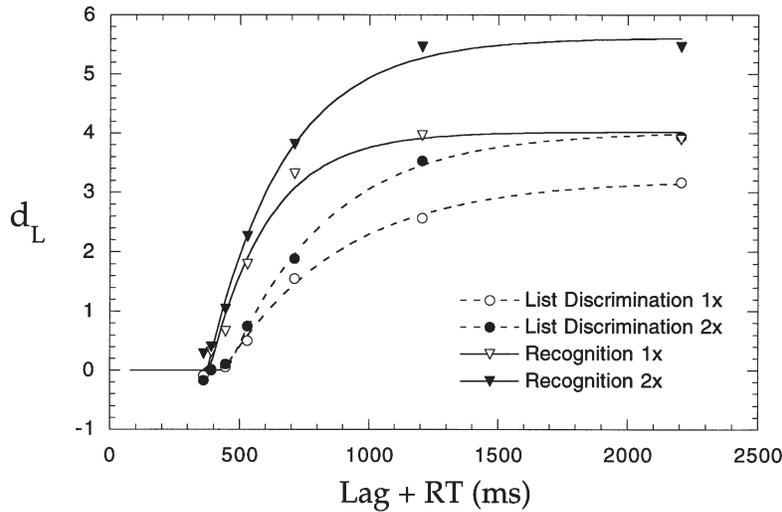
Figure 7. Discriminability curves for the recognition-memory and list-discrimination tasks of Experiment 2. 1×, one presentation; 2×, two presentations.

GENERAL DISCUSSION

These experiments explored the retrieval of information useful for two types of memory judgment. Specifically, it examined the time course of retrieval for list-membership decisions and compared it with that for old–new recognition judgments. Earlier results using the response-signal method showed that the retrieval intercept for recognition memory precedes that for recalling whether a noun was studied in singular or plural form (Hintzman & Curran, 1994), and also precedes the intercept for recalling a word's auditory or visual presentation modality (Hintzman & Caulton, 1997). The latter study found, in addition, that repeating a word three times in the same modality changed the retrieval dynamics of the modality-judgment task, shortening the intercept and prolonging the approach to asymptote. No such effects of repetition were observed in the recognition-memory task. On the basis of the working hypothesis that there are two retrieval processes in episodic memory—a fast familiarity process and a slower process underlying recall of an event's content and context—we predicted that the time course of list discrimination would be similar to those of plurality and modality information in the previous studies.

In that context, the present experiments appear to present a simple picture. In Experiment 1, which used an exclusion design, the initial retrieval of familiarity (the intercept for discriminating both lists from neither list) occurred 115 ± 46 msec earlier than retrieval of information on list membership (the intercept for discriminating the target list only from the nontarget list only). In Experiment 2, which used separate recognition and list-discrimination instructions, the difference was 104 ± 32 msec at frequency = 1 and 94 ± 29 msec at frequency = 2. The differences obtained in Experiments 1 and 2 are statistically indistinguishable, and are well within the range the earlier studies would lead one to expect.[4]

It is interesting that the estimates from Experiments 1 and 2 were so much alike. We had speculated that the exclusion task of Experiment 1 might have artifactually delayed list discrimination, because subjects had the dual task of rejecting both new words and familiar words from

These data are consistent with the view that minimal retrieval times primarily reflect task differences. Although repetition affected asymptotic performance in both the recognition and list-discrimination tasks, it had little or no effect on the intercepts.

Table 2
Summary Statistics for Fits of Shifted
Exponential Function to Data of Experiment 2

| Curve | | Parameter | | | | | |
| | | I (msec) | | $1/R$ (msec) | | A | | |
Task	Frequency	M	SE	M	SE	M	SE	Mean r^2
Recognition	1	402	15	206	27	3.99	0.36	0.86
Recognition	2	388	10	240	17	5.32	0.44	0.89
List discrimination	1	506	27	235	42	2.96	0.33	0.79
List discrimination	2	481	28	306	51	3.92	0.48	0.81

Note—I, intercept; R, rate; A, asymptote; r^2, proportion of variance explained.

the nontarget list. This would not have been a factor in Experiment 2, where recognition and list discrimination were performed separately. There are, of course, other differences between the two experiments that confound direct comparisons of the results, but the similarity of the intercept differences suggests that the complexity of the task in Experiment 1 did not appreciably delay the list-discrimination intercept. This result reinforces a recent conclusion of Buchner, Erdfelder, Steffens, and Martensen (1997)—that parameters of the exclusion task can be recovered from a source-monitoring paradigm—and extends it to retrieval dynamics.

What can we conclude about effects of repetition? As mentioned earlier, Hintzman and Caulton (1997) found that increasing the number of presentations from 1 to 3 shortened the intercept and slowed the rate of approach to asymptote in the modality-judgment task, while no such tendencies were seen in the case of recognition memory. Similar interactions between task and frequency were not observed in the present Experiment 2. However, examination of Table 2 shows that the list-discrimination parameter estimates did change in the right directions. At frequency = 2, the intercept was 25 $\pm$ 29 msec shorter than at frequency = 1, and the inverse of the rate $(1/R)$ was 71 $\pm$ 51 msec longer. Neither of these differences was statistically significant, but both were in the direction to be expected on the basis of the earlier modality-judgment data. Since a frequency manipulation of 3 versus 1 should produce larger effects than one of 2 versus 1, and the differences are in the predicted directions, the present results have to be considered ambiguous. They do not support the inference that repetition affects the retrieval dynamics of list discrimination, but neither are they inconsistent with that hypothesis.

These experiments are not the first to use the list-discrimination task to separate hypothetical processes of familiarity and recall, and so it is worth considering how they differ from the previous studies. Yonelinas (1994) and Yonelinas and Jacoby (1994) applied the process-dissociation method (Jacoby, Toth, & Yonelinas, 1993) to a list-exclusion task, in an attempt to obtain separate estimates of the contributions of familiarity and recall (recollection) to recognition memory. Two of the assumptions underlying the model that they used deserve mention. One is that recall and familiarity are independent—that is, for a given subject, the probability of recall is no higher for words that seem familiar than it is for words that seem unfamiliar. (For an extended debate on the degree to which violation of this assumption could distort results, see Curran & Hintzman, 1995, 1997; Hintzman & Curran, 1997b; Jacoby, Begg, & Toth, 1997; Jacoby & Shrout, 1997). The other assumption is that if a word's study episode can be recalled, list discrimination will be perfect. That is, as the process-dissociation equations are applied by Yonelinas and by Yonelinas and Jacoby, all correct recognition above that measured by correct list identification is attributed to familiarity.

In contrast, we do not assume that the asymptote of list discrimination represents the asymptotic level of recall.

One reason we think they are different is that subjects can often recall whether a word was nearest the beginning, middle, or end of a list even when they cannot remember which list it was in (Hintzman, Block, & Summers, 1973). It seems likely that memory for a word's input modality would similarly be above chance even when list membership is misremembered. In accordance with Johnson's source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), we assume that subjects can often recall many things about a word's earlier presentation that are of little or no use in list discrimination. Some recalled information may be useful in list discrimination, some in judging study modality, and some in judging the word's within-list position. Rather than relying on the process dissociation equations, our strategy uses the response-signal method to separate familiarity and recall on the basis of the estimated point in the retrieval episode when task-relevant information first influences decisions. We maintain the working hypothesis that there are just two basic retrieval processes because, in the studies we have done so far, intercepts for recall of plurality, judgments of modality, and judgments of list membership have been delayed, relative to the intercept for recognition memory, by about the same amount.

This raises the question of what kind of information underlies the ability to identify the list in which a test word was studied. Since lists necessarily differ in recency, and familiarity may decay over time, it seems natural to assume that subjects will try to base list discrimination on familiarity. Certain aspects of our data suggest that they do this: In Experiment 1, although the difference was not significant, subjects seem to have relied more on familiarity when the target was List 2 than when it was List 1. And, in the list-discrimination task of Experiment 2, subjects tended to assign nonrepeated words to List 1 and repeated words to List 2. Familiarity would have been an unreliable clue to list membership in either experiment, however, because some words were presented twice and others were presented once. The retrieval intercepts of Experiments 1 and 2 show that information that is available early, and is sufficient for above-chance recognition, is not sufficient for correctly identifying the list in which the test word was studied. It is clear therefore that recency is not the only basis of list-discrimination performance. Other factors have been labeled—vaguely—as context (Hintzman et al., 1973) and organization (Hintzman & Waters, 1970). In the present experiments, of course, subjects knew that a list-discrimination test was likely, and had plenty of opportunity to develop a strategy for dealing specifically with this task. We did not collect strategy reports from our subjects. However, subjects in memory experiments sometimes volunteer the information that they made up stories connecting to-be-remembered words to an overall schema. That strategy would seem ideal for a list-discrimination experiment. If one could recall a test word's story context and the order in which the stories were constructed, then judging the word's list membership would be easy.

If the distinction between familiarity and recall processes is valid, this has implications for theories of recognition

memory. The most common approach to recognition has been to assume that old and new test items differ along a continuum of strength or familiarity. Anderson and Bower (1972) explicitly rejected this approach, however, on the basis of the demonstrated abilities of subjects to remember such aspects of a word's presentation as its list membership, frequency, recency, and spacing of repetitions. They argued that strength was inadequate to explain such feats because it was "ahistorical," in that a single measure reflected the test item's past frequency, recency, and exposure duration. Anderson and Bower (1972) proposed that familiarity theory be replaced by one in which recognition judgments were based on the retrieval of associations to context—essentially, the process of recall or recollection. On this view, recognition and list-discrimination judgments are not seen as fundamentally different: Recognition requires discriminating the recalled experimental context from nonexperimental contexts, just as list discrimination requires discriminating the recalled context of List 1 from that of List 2.

It seems undeniable that recognition memory has this character, because recognition instructions always refer, either explicitly or implicitly, to past occurrences of items in a particular context or list. In the vast majority of recognition experiments, however, the relevant context is confounded with recency, because a relatively short time separates study and test. Under these circumstances, if presentations of an item were to endow it with familiarity that declines over time, this "ahistorical" cue would be highly relevant to the decision involving experimental context. A sensible subject could be expected to make use of this redundant cue—especially if familiarity were a simple, easy-to-evaluate quantity that becomes available earlier in retrieval than does the recall of context. Of course, our data may not demand such an interpretation. A context-retrieval theorist could assume that discriminating between the List 1 and List 2 contexts requires a more fine-grained analysis than does discriminating between experimental and extra-experimental contexts. The trick would be to explain why this difference in contextual detail yields a difference in the intercepts of the response-signal curves for the recognition and list-discrimination tasks and not just a difference in growth rates and asymptotes.

This problem is not confined to a specific framework or theory, however, because few current approaches to memory directly address questions of retrieval dynamics. Perhaps the most obvious interpretation of the observed intercept difference is that list discrimination involves an additional processing stage that adds about 100 msec or so to the time required for recognition. The memory literature suggests three general ways in which such a stage might be characterized. One is to say that recognition judgments can be made with little (or no) reflection, while conscious mediation is required for list-discrimination judgments (e.g., Jacoby, 1991; Johnson et al., 1993). According to one empirical estimate, it can take as long as 500 msec to become consciously aware of a stimulus (Libet,

1993). This would more than account for a 100-msec difference in intercepts.

A second way to characterize the observed intercept difference is in terms of brain mechanisms. As mentioned earlier, it has been proposed that "parahippocampal" regions of the medial temporal lobe can preserve the intermediate-term familiarity of a stimulus, while the hippocampal formation encodes relations among stimuli (Eichenbaum et al., 1994). Internal and external connections of the hippocampus appear appropriate for such associative learning, and for reactivating cortical areas that were involved in the original experience (Heil, Rösler, & Hennighausen, 1996; Marr, 1971; Treves & Rolls, 1993). Thus, hippocampal structures may be required for recall of associative information, such as that related to list membership, but not for recognition on the basis of familiarity (see Aggleton & Shaw, 1996). The frontal lobes may also be implicated in memory search and in the interpretation of associatively retrieved information (e.g., Janowsky, Shimamura, & Squire, 1989; Milner, Corsi, & Leonard, 1991; Moscovitch, 1992). The involvement of such additional neural circuitry could help explain why the intercept for list discrimination is delayed relative to that for recognition.

A third way to characterize the task difference is in terms of a computational model. In "global matching" models, recognition judgments are based on a simple, unidimensional strength or familiarity scale, which reflects the degree to which the test item matches memory as a whole. These models suggest various processes, any of which could add complexity—and therefore time—in a task requiring discrimination of list membership. List discrimination might require assembling or computing a more complex retrieval cue, searching memory sequentially, and identifying or evaluating information that has been retrieved (for a general discussion of global matching models, see Clark & Gronlund, 1996). Explanations of retrieval-dynamics differences that invoke consciousness or brain mechanism do not, of course, obviate the need for a computational model. It is still necessary to specify what processes conscious awareness, or the different brain mechanisms, perform.

REFERENCES

Aggleton, J. P., & Shaw, C. (1996). Amnesia and recognition memory: A reanalysis of psychometric data. *Neuropsychologia*, **34**, 51-62.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, **79**, 97-123.

Anderson, J. R., & Bower, G. H. (1974). Interference in memory for multiple contexts. *Memory & Cognition*, **2**, 509-514.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583-612). New York: Academic Press.

Barnett, V., & Lewis, T. (1984). *Outliers in statistical data* (2nd ed.). Chichester, U.K.: Wiley.

Buchner, A., Erdfelder, E., Steffens, M. C., & Martensen, H. (1997). The nature of memory processes underlying recognition judgments in the process dissociation procedure. *Memory & Cognition*, **25**, 508-517.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of

recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, **3**, 37-60.

Corbett, A. T., & Wickelgren, W. A. (1978). Semantic memory retrieval: Analysis by speed accuracy tradeoff functions. *Quarterly Journal of Experimental Psychology*, **30**, 1-15.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 531-547.

Curran, T., & Hintzman, D. L. (1997). Consequences and causes of correlations in process dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 496-504.

Dosher, B. A. (1981). The effects of delay and interference: A speed-accuracy study. *Cognitive Psychology*, **13**, 551-582.

Dosher, B. A. (1984a). Degree of learning and retrieval speed: Study time and multiple exposures. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 541-574.

Dosher, B. A. (1984b). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology*, **16**, 519-555.

Dosher, B. A., & Rosedale, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory & Language*, **30**, 125-160.

Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.

Eichenbaum, H., Otto, T., & Cohen, N. J. (1994). Two functional components of the hippocampal memory system. *Behavioral & Brain Sciences*, **17**, 449-518.

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

Gronlund, S. D., Edwards, M. B., & Ohrt, D. D. (1997). Comparison of the retrieval of item versus spatial position information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1261-1274.

Gronlund, S. D., & Ratcliff, R. (1989). The time-course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 846-858.

Heil, M., Rösler, F., & Hennighausen, E. (1996). Topographically distinct cortical activation in episodic long-term memory: The retrieval of spatial versus verbal information. *Memory & Cognition*, **24**, 777-795.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528-551.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, **88**, 297-306.

Hintzman, D. L., Block, R. A., & Summers, J. J. (1973). Contextual associations and memory for serial position. *Journal of Experimental Psychology*, **97**, 220-229.

Hintzman, D. L., & Caulton, D. A. (1997). Recognition memory and modality judgments: A comparison of retrieval dynamics. *Journal of Memory & Language*, **37**, 1-23.

Hintzman, D. L., Caulton, D. A., & Curran, T. (1994). Retrieval constraints and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 275-289.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.

Hintzman, D. L., & Curran, T. (1997a). Comparing retrieval dynamics in recognition memory and lexical decision. *Journal of Experimental Psychology: General*, **126**, 228-247.

Hintzman, D. L., & Curran, T. (1997b). More than one way to violate independence: Reply to Jacoby and Shrout. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 511-513.

Hintzman, D. L., & Waters, R. M. (1970). Recency and frequency as factors in list discrimination. *Journal of Verbal Learning & Verbal Behavior*, **9**, 218-221.

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

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

Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the process-dissociation procedure? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 484-495.

Jacoby, L. L., & Shrout, P. E. (1997). Toward a psychometric analysis of violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 505-510.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, **122**, 139-154.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, **27**, 1043-1056.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.

Libet, B. (1993). *Neurophysiology of consciousness: Selected papers and new essays of Benjamin Libet*. Boston: Birkhauser.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of "nonparametric" measures of sensitivity and response bias. *Psychonomic Bulletin & Review*, **3**, 164-170.

Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical transactions of the Royal Society of London*, **262**, 23-81.

McElree, B., & Dosher, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, **122**, 291-315.

Milner, B., Corsi, P., & Leonard, G. (1991). Frontal-lobe contribution to recency judgments. *Neuropsychologia*, **29**, 601-618.

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, **4**, 257-267.

Mulligan, N., & Hirshman, E. (1995). Speed-accuracy trade-offs and the dual process model of recognition memory. *Journal of Memory & Language*, **34**, 1-18.

Murdock, B. B., Jr. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89**, 609-626.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words* (Progress Rep. No. 10). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1992). Models for recall and recognition. *Annual Review of Psychology*, **43**, 205-234.

Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology*, **21**, 139-155.

Reed, A. V. (1973). Speed–accuracy tradeoff in recognition memory. *Science*, **181**, 574-576.

Reed, A. V. (1976). List length and the time-course of recognition in immediate memory. *Memory & Cognition*, **4**, 16-30.

Rensink, R. A. (1990). Toolbox-based routines for Macintosh timing and display. *Behavior Research Methods, Instruments, & Computers*, **22**, 105-107.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Application to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

Treves, A., & Rolls, E. T. (1993). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, **4**, 374-391.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, **41**, 67-85.

Wickelgren, W. A. (1981). Human learning and memory. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 32, pp. 21-52). Palo Alto, CA: Annual Reviews.

Wickelgren, W. A., & Corbett, A. T. (1977). Associative interference

and retrieval dynamics in yes–no recall and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **3**, 189-202.

WINOGRAD, E. (1968). List differentiation as a function of frequency and retention interval. *Journal of Experimental Psychology Monographs*, **76**, (2, Pt. 2).

YONELINAS, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1341-1354.

YONELINAS, A. P., & JACOBY, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, **48**, 516-534.

NOTES

1. An idiosyncrasy in recent versions of the Macintosh Sound Manager, unrecognized at the time these experiments were conducted, delays the onset of a sound by an amount that varies depending on the configuration of the system and machine. For this reason, the stated lags probably underestimate the true lags by tens of milliseconds. The onset delay confounds comparisons of absolute times—especially of retrieval intercepts—across experiments. However, it does not affect within-experiment comparisons such as those on which the conclusions of this article are based.

2. This is based on the reduced sample of $N = 21$, with outliers removed. See Hintzman and Curran (1994), Table 2.

3. Readers accustomed to choice-RT data may be surprised that such small differences are statistically significant. It is worth noting that long-latency responses ($RT > 350$ msec) were trimmed from the data, and that the response-signal paradigm gives most control of processing time to the experimenter. Small RT error variances reflect the extent of that control.

4. Hintzman and Curran (1997a, Figure 12) compare recognition intercepts over several published experiments from our laboratory. The recognition intercepts obtained here are generally 20–60 msec shorter than those earlier ones. For the probable explanation of this discrepancy, see note 1.

A recent study compared the retrieval dynamics of recognition memory with those of judgments of spatial location (Gronlund, Edwards, & Ohrt, 1997). When subjects simultaneously had to reject new items and old items that were in the wrong location, the intercept for old versus new averaged 122 msec (Experiment 1) and 61 msec (Experiment 2) earlier than that for correct versus incorrect location. When test blocks for old versus new and location judgments were separated, the intercept for old versus new averaged 122 msec earlier than that for location (Experiment 3). These intercept differences are consistent with ours, both in magnitude and in apparent insensitivity to combined versus separate task instructions.

APPENDIX
Data Deleted From Experiment 2

Table A1 summarizes the curve-fitting statistics for the 10 subjects whose data were deleted from Experiment 2 because of low r^2 and outliers. In comparison with the retained subjects (Table 2), asymptotes tended to be lower, especially in list discrimination, and intercept estimates for list discrimination were longer. Low r^2 and low asymptotes make it difficult for a curve-fitting algorithm to discriminate between chance and above-chance performance. This may be why the list-discrimination intercept estimates for the deleted subjects are noticeably longer than the ones in Table 2.

ANOVAs on the data summarized in Table A1 below were nevertheless consistent with those for the retained subjects. Intercepts were reliably affected by task [$F(1,9) = 12.19$, $MS_e = 48,459$], and asymptotes were affected by task [$F(1,9) = 23.05$, $MS_e = 1.18$] and by repetition [$F(1,9) = 16.63$, $MS_e = 0.69$]. No other effects were significant.

Table A1
Summary Statistics for Curve Fits of the 10 Subjects Deleted From Experiment 2

| Curve | | Parameter | | | | | | |
| | | I (msec) | | $1/R$ (msec) | | A | | |
Task	Frequency	M	SE	M	SE	M	SE	Mean r^2
Recognition	1	416	57	237	39	3.40	0.32	0.77
Recognition	2	387	25	286	64	4.63	0.44	0.86
List discrimination	1	599	107	175	55	1.91	0.45	0.51
List discrimination	2	690	85	266	69	2.83	0.66	0.56

Note—I, intercept; R, rate; A, asymptote; r^2, proportion of variance explained.