Journal of Experimental Psychology:
Learning, Memory, and Cognition
1993, Vol. 19, No. 5, 1024–1039

0278-7393/93/$3.00

Proactive Interference and the Dynamics of Free Recall

John T. Wixted and Doug Rohrer

Proactive interference (PI) has long been recognized as a major cause of forgetting. We conducted two experiments that offer another look at the subject by providing a detailed analysis of recall latency distributions during the buildup of and release from PI. These functions were accurately characterized by the convolution of the normal and exponential distributions (viz., the ex-Gaussian), which previously has been shown to describe recognition latency distributions. Further, the fits revealed that the increase in recall latency associated with the buildup of PI results from a slowing of the exponential retrieval stage only. The same result was found even when a short retention interval was used (and recall probability remained constant). These findings suggest that free-recall latency may be a sensitive index of the increased search set size that has often been assumed to accompany the buildup of PI.

A central insight emerging from the memory literature of the 1950s and 1960s was that previously learned information can result in the rapid forgetting of more recently learned information. Underwood (1957) argued that this phenomenon, termed *proactive interference* (PI), was by far the major cause of forgetting in everyday life. Indeed, even in laboratory experiments, the degree of retroactive interference encountered over the course of hours or days was assumed to pale in comparison with the degree of proactive interference resulting from years of prior learning. Although its preeminent (and still unexplained) role in the process of forgetting continues to be recognized, interest in the subject of PI has waned in recent years. The present article contributes a new empirical analysis of this important subject and pursues a detailed theoretical exploration into its underlying nature.

In a typical PI experiment, subjects receive blocks of Brown–Peterson trials involving words from a single category (Wickens, 1972). Within a block, free-recall performance declines with each successive trial (the buildup of PI) but recovers each time a new category is introduced (release from PI). In most cases, the dependent variable used in these experiments was the percentage of correct free-recall responses. However, in the research to be presented here, we focus on latency to free recall. Research on free-recall latency in any context is very limited, and in the study of PI it is almost nonexistent.

Why might free-recall latency be an interesting variable to investigate? Because such a measure provides important information about the process of retrieval that is likely to be missed by static measures, such as probability of recall. Before addressing the question of exactly what that information might be, we review the scant literature pertaining to the more general and purely empirical question of whether these two measures of memory performance are truly independent. If probability and latency measures always covary, then they would simply represent redundant rather than independent measures of memory.

Probability and Latency of Free Recall

In the 1950s, Bousfield and his colleagues investigated the viability of Marbe's law, which basically states that items associated with a high average probability of recall (e.g., primacy and recency items) will also tend to be recalled prior to other items (i.e., they will have a relatively short recall latency). Several experiments involving retrieval from semantic or episodic memory supported this intuitively appealing idea (Bousfield & Barclay, 1950; Bousfield, Cohen, & Silva, 1956; Bousfield, Whitmarsh, & Esterson, 1958). For example, Bousfield et al. (1958) reported that for lists of 10 words or more, an item's output position in immediate free recall was inversely related to its overall probability of recall.

Note that Marbe's law specifies which items within a set will be recalled prior to the others, but it says nothing about the absolute time required for retrieval. Whether overall retrieval time involves seconds, minutes, or hours, Marbe's law simply asserts that items associated with a higher probability of recall will be the ones that are retrieved first. Thus, the law does not necessarily imply that the average latency to recall associated with a set of items will decrease as the average probability of recall associated with those items increases. Indeed, the distinction between relative latency (i.e., the order in which items are retrieved) and absolute latency (i.e., the average time to retrieval) is of critical importance from both an empirical and a theoretical point of view.

One way to illustrate how these two latency measures might behave in different ways is by means of a hypothetical retrieval model that consists of two assumptions: (a) Items are randomly sampled from a search set and are then replaced and (b) the more rehearsal an item receives, the more copies of that item reside in the search set. The first assumption regarding the nature of retrieval has been repeatedly suggested over the years, and it served as the original framework

John T. Wixted and Doug Rohrer, Department of Psychology, University of California, San Diego.

Correspondence concerning this article should be addressed to John T. Wixted, Department of Psychology 0109, University of California, San Diego, La Jolla, California 92093.

of the search of associative memory (SAM) model for recall (Rundus, 1973; Shiffrin, 1970). The second assumption represents the essence of Bernbach's (1970) multiple copy model of human memory, and it forms one component of Laming's (1992) theory of retention on the Brown–Peterson task. In a typical list learning experiment, the multiple copy model assumes that primacy items are multiply represented within the search set because they receive the most rehearsal.

What does this hypothetical model predict about order of recall and average latency to recall? Under poor learning conditions (e.g., a fast rate of presentation), only a few of the list items will be encoded in the search set, which means that the probability of recall will be low. Under more favorable learning conditions, more list items will be stored in the search set and the probability of recall will be higher. In both conditions, primacy items are likely to be multiply represented (because those items receive the most rehearsal) and thus are likely to be sampled and recalled first. Therefore, in accordance with Marbe's law, items with the highest probability of recall will be output first in both conditions. However, average search time will actually increase (not decrease) under more favorable learning conditions because of the extra time required to sample (and resample) the additional items contained within the search set.

Although this hypothetical model illustrates why Marbe's law may not apply to absolute recall latency, the question is really an empirical one. Very few studies have attempted to evaluate the effect of experimental manipulations on average latency to free recall, but, again, Bousfield provides some of the relevant data. Bousfield, Sedgewick, and Cohen (1954) performed a simple experiment in which subjects were presented with a list of 60 words to memorize. One group of subjects received a single presentation of the list, whereas other groups received two, three, four, or five presentations prior to recall. Obviously, repeated presentations served to increase the probability of recall, but the question of interest here concerns the effect on average latency to recall.

Bousfield et al. (1954) did not actually report latency data but instead reported the rate of approach to the cumulative recall asymptote. The results of their analysis indicated that as the probability of recall increased, the rate of approach decreased. This finding is relevant to the present discussion because a slower rate of approach to asymptote implies a longer average recall latency. Table 1 provides the obtained probabilities of recall as well as the rate parameter obtained from the best-fitting hyperbola in each condition. The rate parameter is equal to the time required (in minutes) to reach half asymptotic recall levels. These data now reveal that as probability of recall increased, the average latency to recall increased as well (Condition 3 being the one exception to this general trend). Note that this result is not actually in conflict with Marbe's law; within each condition, individual items with the highest probability of recall were probably the ones that were retrieved first. Nevertheless, the average time to recall increased as subjects encoded more of the list items.

In their summary of the limited research on this issue, Roediger, Stellon, and Tulving (1977) stated that the relationship reported by Bousfield et al. (1954) could be regarded as the standard finding. Thus, perhaps it would be reasonable to suppose that whenever more items are recalled, average latency to recall will increase. However, Roediger et al. reported an exception to this general rule in the context of part-list cuing. In that experiment, part-list cues served to decrease the probability of recall and to increase latency of recall.

Table 1
Recall Probability and Rate Parameter Measures From Bousfield, Sedgewick, and Cohen (1954)

	Recall measures	
List presentations	Probability	Rate parameter
1	.40	0.0235
2	.49	0.0164
3	.53	0.0191
4	.59	0.0127
5	.63	0.0108

Note. From "Certain Temporal Characteristics of the Recall of Verbal Associates" by W. A. Bousfield, C. H. W. Sedgewick, and B. H. Cohen, 1954, *American Journal of Psychology, 67,* p. 115. Copyright 1954 by the University of Illinois Press. Adapted by permission.

Although these results sound contradictory, that may not be the case. Indeed, it may turn out that no a priori empirical relationship between average probability of recall and average latency to recall can be specified because these two dependent variables reflect different memory processes. The idea that probability and latency measures ought to be related is based on the implicit assumption that both measures reflect some unitary property of memory (e.g., trace strength). A detailed and informative analysis of this idea as it relates to cued-recall latency was presented by MacLeod and Nelson (1984). After reviewing the literature and presenting three experiments of their own, MacLeod and Nelson concluded that probability of recall and latency to recall were sometimes positively correlated, sometimes negatively correlated, and sometimes altogether uncorrelated. On the basis of these findings, they rejected the idea of unidimensionality and concluded that these two measures capture different aspects of memory. For reasons that are made clear later, we reach the same conclusion with respect to free-recall latency: Probability of recall reflects item accessibility, whereas latency reflects (among other things) the breadth of mental search. Although this account is at odds with a unidimensional view, it is perfectly compatible with most latency analyses involving recognition memory (e.g., Atkinson & Juola, 1974).

Theories of PI

Although, apparently, predictions about the effects of PI on free-recall latency cannot be derived from any known empirical law, the same cannot be said of current theories of PI. According to the most widely accepted account, the buildup of PI reflects a growing impairment in the ability to distinguish items that appeared on the most recent list from those that appeared on earlier lists (Baddeley, 1990; Crowder, 1976; Underwood, 1945). This temporal discrimination theory explicitly assumes that subjects are unable to restrict their search to the most recent list of items and instead search the entire set of category-specific items that have been presented thus far.

If that account is true, what should happen to free-recall latency? On Trial 1, subjects must search through a relatively small number of representations and so should find each one quickly. On Trial 2, they must search through the items presented on the first trial as well as those presented on the second. As a result, the recall latency for the items presented on Trial 2 should increase owing to the extra search time. On Trial 3, latency should increase even further as subjects search through items from all three trials.

Although the temporal discrimination account is probably the most widely accepted theory of PI, a series of experiments reported by Dillon and his associates produced results that are very difficult to reconcile with that analysis (Dillon, 1973; Dillon & Bittner, 1975; Dillon & Thomas, 1975). If the central difficulty is one of distinguishing between current and preceding list items, then performance on later trials should improve considerably if (a) subjects are provided with the earlier (and now incorrect) items prior to recall or (b) subjects are permitted to recall all of the category-relevant items they can remember on each trial. However, contrary to the temporal discrimination account, neither manipulation appreciably improved recall on later trials. Moreover, Dillon and his colleagues found that when intrusions did occur, subjects had little difficulty identifying which of their own responses were correct and which were words from previous trials. Dillon thus concluded that the buildup of PI does not reflect a temporal discrimination problem so much as it reflects an inability to access the correct representation in the first place (i.e., search failure). If so, then the size of the mental search set (and the time taken to search through it) might not be expected to increase with the buildup of PI.

Although temporal discrimination theory clearly predicts a longer search time on later trials, the simple demonstration that mean recall or recognition latency increases with PI is actually not very informative (Anderson, 1981; Gorfein & Jacobson, 1973). For example, many theories of memory hold that retrieval involves at least two stages: The first stage involves establishing a relatively narrow search set (e.g., via self-generated retrieval cues), and the second stage involves sampling the items within that set (e.g., Glenberg & Swanson, 1986; Rundus, 1973; Shiffrin, 1970). It is possible that the buildup of PI delays the first stage (i.e., search onset) without affecting the second, in which case there would be no reason to assume that the size of the search set itself increases. Whether PI affects search onset, search time, or both cannot be determined using a simple summary measure such as mean latency to recall. The only way to distinguish between possibilities such as these is to examine free-recall latency distributions. Indeed, the importance of studying reaction time (RT) distributions in favor of simple summary statistics has been made repeatedly in the context of recognition memory (e.g., Heathcote, Popiel, & Mewhort, 1991; Hockley, 1984; Ratcliff & Murdock, 1976). The same arguments apply here as well.

Free-Recall Latency Distributions

The preceding discussion reveals that even if it had been possible to identify an empirical law relating mean probability of recall to mean latency, such a law would be incomplete because changes in mean latency can occur in a variety of ways. Free-recall latency distributions, which help to distinguish between the various possibilities, can be plotted in either cumulative or noncumulative form. The cumulative distribution shows the total number of items recalled up to each point in the recall period and can be roughly described as a negatively accelerated function that rises from zero to some finite asymptote. With a few exceptions (e.g., Bousfield et al., 1954), most of the past research on this subject suggests that the function is reasonably well described by an exponential of the form $F(t) = N(1 - e^{-t/\tau})$, where N represents the number of items recalled given unlimited time (i.e., N represents asymptotic recall) and τ represents the average latency to recall associated with those N items (Bousfield & Sedgewick, 1944; Indow & Togano, 1970; Roediger et al., 1977). For most of our analyses, we have found the corresponding noncumulative distribution to be very informative as well. This distribution shows the number of items recalled during each interval of the recall period and is described by the equation $f(t) = (N/\tau)e^{-t/\tau}$, where the parameters have the same meanings as before. Note that $f(t)$ is simply the first derivative of $F(t)$.

The question of interest here concerns the way in which these distributions might change as mean recall latency changes. Figure 1 shows the two simplest possibilities. The upper panel shows three curves that reflect an increased latency score resulting from a delay in initiating recall. These curves are actually described by the function $F(t) = N(1 - e^{-(t-c)/\tau})$, where c represents the recall onset latency. Note that these curves were produced by reducing the value of N from Trial 1 to Trial 3 (reflecting the buildup of PI) and increasing the values of c from Trial 1 to Trial 3 (reflecting a growing delay in the onset of recall). The value of τ, which actually represents mean latency measured from the onset of recall, was held constant. This pattern of results might be expected if PI lengthened the time required to find (or establish) a focused search set on later trials. The lower panel shows three curves that would be expected if increased recall latency occurred because of a slowing of the exponential process. Here the value of τ increases as N decreases, and c is held constant at 0. As we argue later, this pattern is more consistent with the idea that the buildup of PI is associated with a growing mental search set. For the moment, the important point is that mean latency to recall can change for a variety of reasons, only two of which are depicted in Figure 1.

One additional reason to investigate free-recall latency distributions concerns the information provided by their mathematical form. If these curves are truly exponential, that fact is not irrelevant to the understanding of the nature of retrieval. Although the point will be developed more fully in a later section, exponential recall implies a constant probability of retrieval associated with the individual items

(McGill, 1963). That is, the to-be-retrieved items are becoming no more or less likely to be retrieved during the course of the recall period. If the situation were otherwise, the curves would not be described well by the exponential. Thus, theories that assume that an item's momentary probability of retrieval increases or decreases during the course of the recall period are inconsistent with exponential retrieval curves.

In what follows, we report the results of two experiments on the buildup of and release from PI. The first is patterned after the classic experiment by Gardiner, Craik, and Birtwistle (1972), and the second is patterned after the classic experiment by Keppel and Underwood (1962). The novel feature of these new experiments was the timing of each response throughout the recall period in order to evaluate the effect of PI on the free-recall latency distribution.

Experiment 1

The design of the first experiment was based on the standard release-from-PI paradigm in which subjects receive successive blocks of Brown–Peterson trials, with each block involving words from a single category (Wickens, 1972). Gardiner et al. (1972) altered this procedure in one important respect. After three consecutive trials involving words from one category (e.g., flowers), the fourth trial involved words from a category that differed in a subtle way (e.g., wildflowers). Subjects who were not specifically informed of this shift failed to notice it, and their performance continued to decline relative to the previous trial. By contrast, subjects who were informed of the subtle category shift either at encoding (the before condition) or at retrieval (the after condition) exhibited a robust release from PI.

Experiment 1 was a replication of Gardiner et al.'s (1972) classic study, except that a more comprehensive picture of recall performance was obtained by timing the output of each list item during recall. The question of interest is how the recall latency distribution is affected by the buildup of and release from PI (whether release is affected by a before cue or by an after cue). Only a very few studies have ever used the after-cue procedure, and none of these have examined its effect on recall latency.

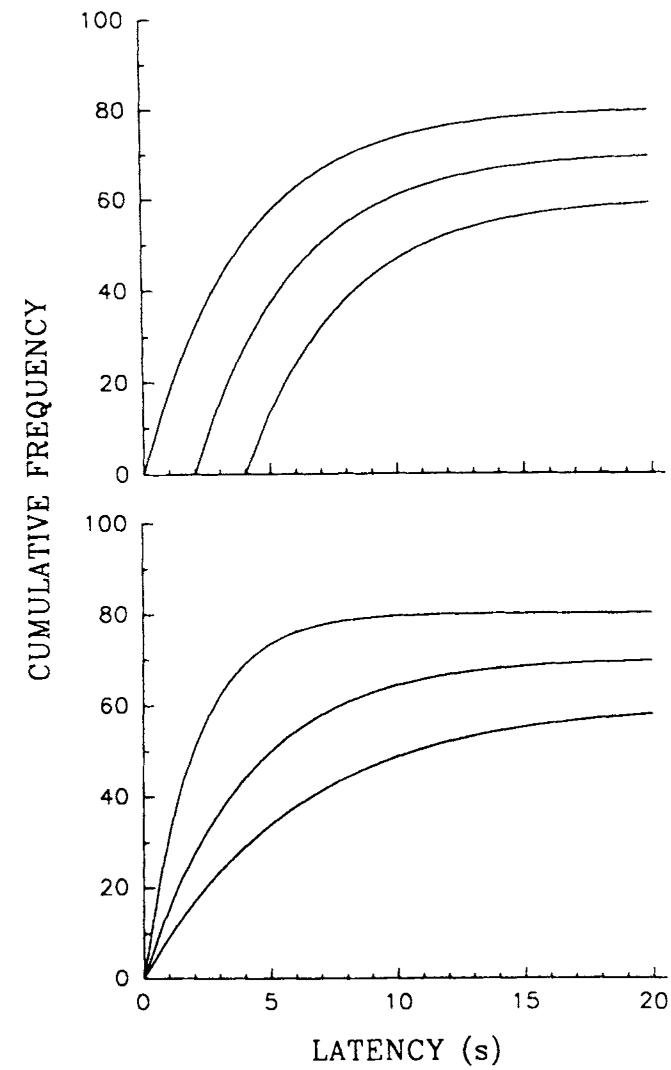

Figure 1. Hypothetical cumulative recall latency distributions reflecting a delay in the onset of recall associated with the buildup of proactive interference (PI; upper panel) and a slowing of an exponential process associated with the buildup of PI (lower panel).

Method

Subjects. The subjects were 36 undergraduates of the University of California, San Diego. Their participation in the experiment satisfied a psychology course requirement.

Materials and design. Three categories of 24 words each were constructed in such a way that half of the items in the category differed in a subtle way from the other half. The categories consisted of body parts (paired vs. singular), states (inland vs. coastal), and sports (indoor vs. outdoor). The items used in the experiment are shown in Table 2. An additional pool of 500 high-frequency words drawn from Thorndike and Lorge (1944) were used as distractors during the retention interval.

The design of the experiment was completely within subjects and involved three conditions: control, before, and after. Each condition consisted of four consecutive Brown–Peterson recall trials involving lists of three items. The items for the first three lists of each condition were randomly drawn without replacement from one subset of a category (e.g., coastal states) whereas those for the fourth list were randomly drawn from the other subset of the same category (e.g., inland states). A different category was in effect for each condition.

The assignment of category to condition and the direction of the subset shift from Trial 3 to Trial 4 (e.g., coastal states to inland states) were counterbalanced across subjects. Subjects were first exposed to one block of four trials without any indication of the subset shift to demonstrate that it was not noticed (the control condition). The next block of four trials was identical except for the presentation of a cue at the beginning of the Trial 4 recall period that described the subset shift (the after condition). The final block of four trials was similar except that the subset cue occurred just prior to the presentation of the fourth list (the before condition).

The before condition always followed the after condition to minimize the possibility that subjects would learn to anticipate the sub-

Table 2
List Materials Used in Experiments 1 and 2

Body parts		States		Sports	
Paired	Singular	Coastal	Inland	Outdoor	Indoor
ankle	chest	Connecticut	Arizona	baseball	aerobics
cheek	chin	Delaware	Colorado	cycling	billiards
ear	face	Florida	Idaho	frisbee	bowling
elbow	forehead	Georgia	Iowa	golf	fencing
eye	heart	Maine	Kansas	lacrosse	gymnastics
heel	mouth	Maryland	Montana	rugby	hockey
knee	navel	Massachusetts	Nebraska	running	ice skating
lip	neck	New York	Nevada	sailing	karate
nostril	nose	Oregon	Oklahoma	skateboarding	ping-pong
palm	spine	Rhode Island	Utah	skiing	racquetball
thumb	throat	Virginia	Wisconsin	surfing	weightlifting
wrist	tongue	Washington	Wyoming	tennis	wrestling

set shift from Trial 3 to Trial 4. If some subjects had received the before condition prior to the after condition, they might have been led to expect a subtle category shift on later trials. We avoided this problem by ensuring that the after condition in the present experiment always preceded the before condition. This arrangement seemed unlikely to introduce any serious interpretive complications. For example, O'Neill, Sutcliffe, and Tulving (1976) previously showed that whether the after cue is presented on the first, second, or third block of four trials, the results are exactly the same. Furthermore, although subjects in Experiment 1 were ultimately informed of the subset shift in the after condition, that knowledge was superfluous in the final set of trials because the before cue was presented prior to the fourth list anyway.

Procedure. All subjects were tested individually and were informed that the purpose of the experiment was to test their memory for word trials. They were further informed that hints would occasionally be provided regarding the list about to be studied or the list about to be recalled. Subjects were explicitly encouraged to use the cue throughout the recall period to remember the list items.

After the instructions, subjects completed 6 practice trials and 12 experimental trials. A trial consisted of the presentation of a list of three words on a computer screen, a 27-s retention interval (involving 18 distractors), and a 20-s oral recall period. Two of the practice trials contained an after cue, 2 contained a before cue, and 2 were presented without any cues. Each practice trial involved words from a different category.

Immediately following the 6 practice trials, the 12 experimental trials were presented. Prior to the presentation of each list, the screen displayed the word *Hint:*, which was followed by the word *none* or, on the fourth trial of the before condition, by a cue that specified the upcoming category subset (e.g., coastal states). Subjects were instructed to read the cue (or the word *none*) aloud to ensure that they attended to it. The hint remained on the screen for 4 s, after which the three list items were presented, one at a time, at a rate of approximately two words per second. Subjects were instructed to read each word aloud as it appeared on the screen. One-half second after the last item was presented, the 27-s retention interval began. The distractor task that was in effect during the retention interval required subjects to decipher and repeat aloud words displayed on the screen in backward order. Thus, for example, if *riahc* appeared, subjects were required to respond by saying *chair.* Eighteen backward distractors were randomly chosen from the high-frequency word pool and were displayed for 1.5 s each during the retention interval.

After the retention interval, the screen again displayed the word *Hint:* along with the word *none* or, on the fourth trial of the after condition, a cue that defined the category subset associated with the previous list of words. Again, subjects were instructed to read the cue (or the word *none*) aloud. Four seconds after the hint appeared, the 20-s oral free-recall period began (signaled by a series of question marks). Between conditions (i.e., after each set of four trials) there was a 90-s rest period.

During the 20-s recall periods, the computer monitor displayed the number of seconds elapsed in the lower center of the screen. Using this information, the experimenter recorded the time at which each word was retrieved by marking a sheet that listed the to-be-recalled items.

Results and Discussion

Figure 2 presents the percentage of correct responses over the four trials of the control, before, and after conditions. As expected, a substantial buildup of PI is evident over the first

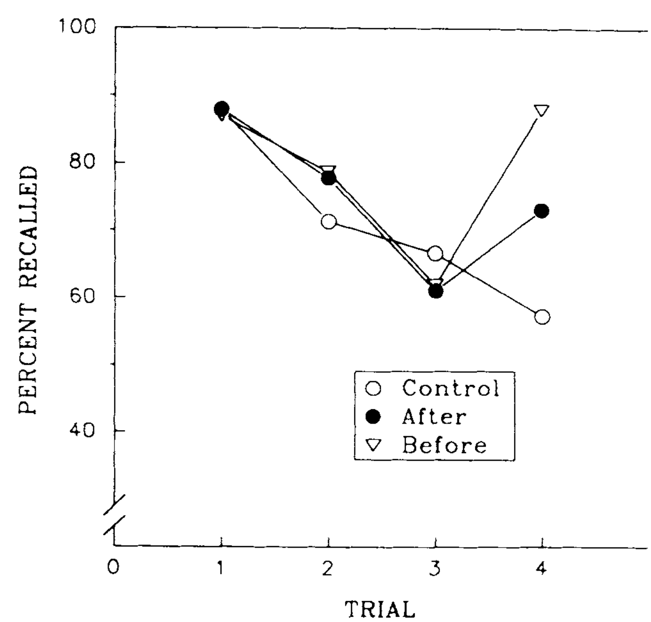

Figure 2. Percentage of items recalled in the control, before, and after conditions of Experiment 1.

three trials of each condition. An analysis of variance (ANOVA) performed on the data from these three trials revealed a highly significant main effect of trial, $F(2, 70) = 26.03$, $MS_e = 14.45$, but the effect of condition and the interaction between trial and condition were not significant (the alpha level was .05 for all statistical tests).

Of more interest was performance on Trial 4 relative to that on Trial 3. The before condition resulted in a robust release from PI, whereas the control condition exhibited a continued buildup of PI. The after condition also exhibited a release from PI, though not quite as large as that obtained in the before condition. To evaluate the significance of these results, we performed 2 × 2 ANOVAs that contrasted the performance in each pair of conditions over Trials 3 and 4. The critical question was whether the interaction between condition and trial was significant. Such a result would suggest that the degree of release in one condition exceeded that in the other. The two analyses contrasting before versus control and after versus control both yielded significant interactions, $F(1, 35) = 15.97$, $MS_e = 10.03$, and $F(1, 35) = 5.23$, $MS_e = 3.67$, respectively. A third test that contrasted the before and after conditions was not significant.

One possible artifactual explanation for the emergence of the after-cue effect in this experiment was that the after cue merely increased the likelihood of guessing correctly. If so, then one might also expect to see a substantial increase in the number of false alarms in the recall protocols. Instead, overt errors (which were only measured for the last 18 subjects) were extremely rare. The numbers of errors relative to the numbers of correct responses on Trials 2–4 are presented in Table 3. Trial 4 errors were negligible in all three conditions, and none were observed in the after condition. On the basis of these results, we conclude that the improvement in performance resulting from the use of before and after cues was probably not the result of guessing.

As we indicated earlier, the principle motivation for this experiment was to evaluate the effects of the buildup of and release from PI on the free-recall latency distribution. This distribution can be plotted either in cumulative form (the usual method) or as a noncumulative function. The principal advantage of the former is that it most clearly reveals any changes in latency owing to a delay in the onset of recall (cf. Figure 1). The main disadvantage is that the data points are not independent, which can create the impression of a smoother retrieval curve than actually exists, and which prohibits the calculation of uncertainty measures associated with parameter estimates. Thus, we begin the analysis by examining the noncumulative functions, and we complete it by analyzing the cumulative function.

Table 3
Overt Intrusions / Correct Responses Observed in Experiment 1

	Condition		
Trial	Control	Before	After
2	2/36	2/43	1/44
3	5/34	1/33	3/33
4	1/31	1/47	0/36

Figure 3 shows the number of items recalled during each 1-s bin of the 20-s recall period. The three graphs in the left panel show the obtained functions for Trials 1–3 averaged across the control, after, and before conditions. Each data point is plotted above the center of its corresponding bin. For example, the first point in each graph is placed above 0.5 s (the center of the bin containing responses emitted between 0 and 1 s), the second above 1.5 s (the center of the bin containing responses emitted between 1 and 2 s), and so on. The three graphs in the right panel of Figure 3 present the Trial 4 data for each of three conditions separately. Because fewer data points are involved, the Trial 4 data are quite a bit more variable.

To evaluate the mathematical form of the curves shown in Figure 3, we fit each data set with the exponential distribution, $f(t) = (1/\tau)e^{-t/\tau}$, where τ reflects the average latency to recall (excluding possible variations in recall onset latency). The fits were accomplished using a maximum likelihood estimation procedure described by Maindonald (1984). The solid curves shown in Figure 3, which represent the best-fitting function for each set of data, reveal that deviations from the exponential are generally nonsystematic. Also shown in the figure are the actual number of items recalled (N), the obtained chi-square values, and the degrees of freedom for each fit. When necessary, cells were combined (and degrees of freedom reduced) to yield an expected value of at least five observations (cf. Hockley, 1984; Ratcliff & Murdock, 1976). None of the chi-square values were significant, which suggests that the fits were adequate. By contrast, fits involving hyperbolic or power functions yielded obvious systematic deviations and significant chi-square values for several conditions. These analyses suggest that retrieval is at least approximately an exponential process and that it remains so after the buildup of PI.

Table 4 presents the obtained values of τ along with their asymptotic standard errors for each of the conditions shown in Figure 3. The standard error estimates are derived from information theory and are calculated from the diagonal elements of the inverse matrix of second partial derivatives evaluated at the optimal parameter values (Wilkes, 1962). These estimates increasingly approximate the true standard errors as sample size becomes large, and they provide one way to gauge the significance of differences between parameter estimates by means of a t test (Ratkowsky, 1983). The values in Table 4 show that the estimated average retrieval latency increased from 2.86 s on Trial 1, to 4.95 s on Trial 2, to 6.61 s on Trial 3. Pairwise t tests revealed that the Trial 2 latency was significantly greater than the Trial 1 latency, $t(21) = 4.80$, and that the Trial 3 latency exceeded that of Trial 2, $t(25) = 2.26$. With regard to release from PI, the τ values shown in Table 4 reveal that the after and before cues presented on Trial 4 resulted in estimated average latencies of 4.72 s and 4.62 s, respectively, which are both shorter than the 6.61-s value observed on Trial 3, $t(20) = 2.14$ and $t(21) = 2.35$, respectively. The Trial 4 recall latency of the control condition (5.35 s) was also unexpectedly shorter than that of Trial 3, but the difference was not significant.

The findings presented in Table 4 suggest that the buildup

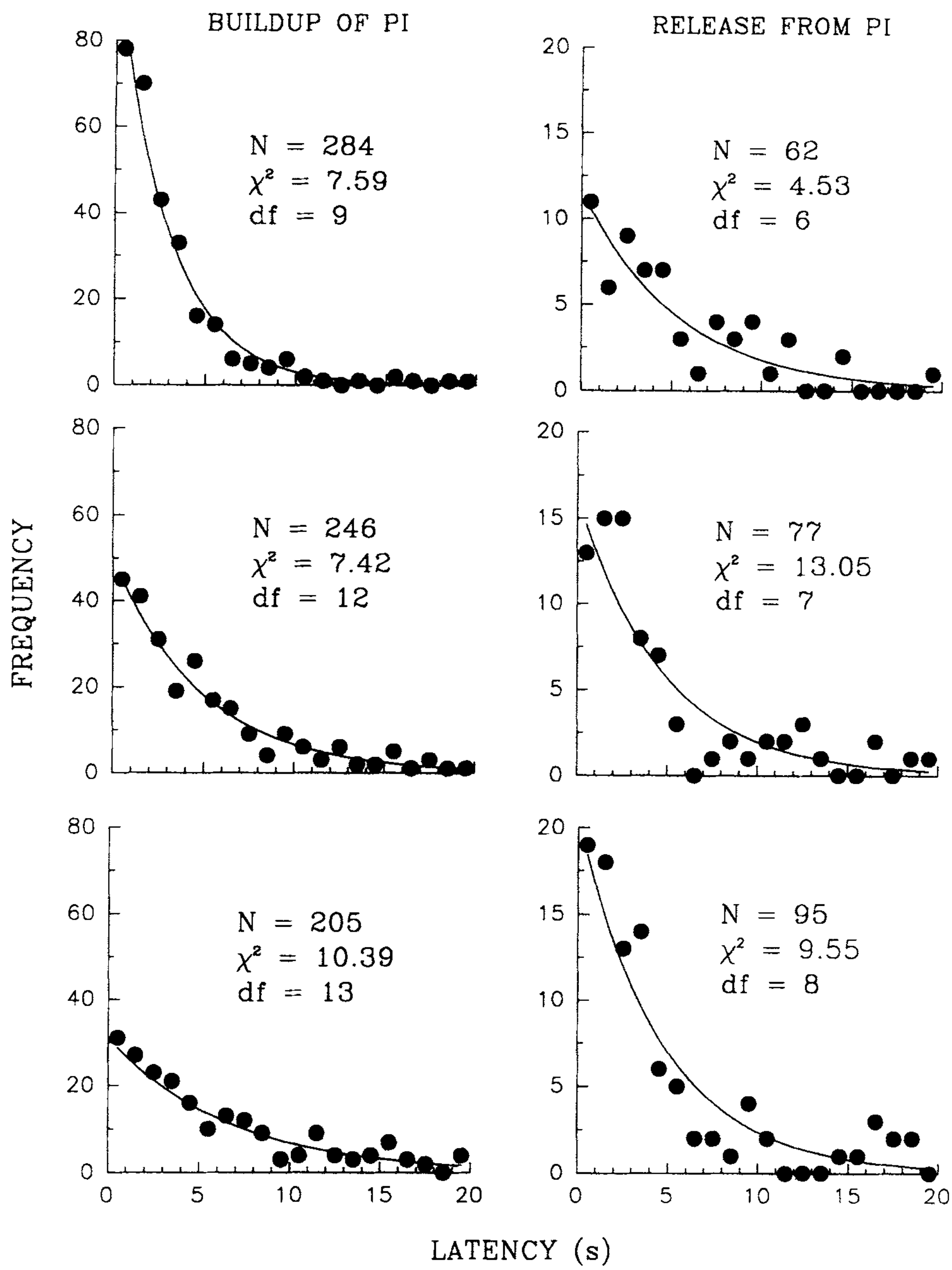

Figure 3. Recall latency distributions averaged over condition for Trials 1–3 (top to bottom, left panel) and recall latency distributions from Trial 4 for the control, after, and before conditions (top to bottom, right panel). (The solid curves represent the best-fitting exponentials. PI = proactive interference.)

of PI results in the slowing of an exponential retrieval process, but there is no evidence that the onset of recall is delayed by PI as well. This latter point is most clearly illustrated by fitting the cumulative latency distributions, $F(t) = 1 - e^{-(t-c)/\tau}$, where c represents a delay in the onset of retrieval. The obtained values of c were all close to zero and were actually slightly negative (–0.49 s, –0.53 s, and –0.61 s for Trials 1–3, respectively). These values suggest that subjects may have covertly initiated recall about one-half second prior to the point at which the recall signal was given. That was possible in this experiment (unlike the next one) because a 4-s interval was interposed between the distractor task and the recall period (time reserved for the presentation of an after cue). Similar values of c were obtained for the control, before, and after conditions on Trial 4 (–0.45 s, –0.74 s, and –0.63 s, respectively).

From an empirical standpoint, the present results indicate that (a) retrieval follows an exponential time course, a fact

Table 4
Exponential Parameter Estimates for the Retrieval Functions From Experiment 1

Trial	τ	Condition	τ
1	2.86 (0.17)	Control	5.35 (0.83)
2	4.95 (0.38)	After	4.72 (0.63)
3	6.61 (0.62)	Before	4.62 (0.56)

Note. The parenthetical values represent the asymptotic standard errors of the parameter estimates. τ = average time of the exponential stage.

that remains true even after the buildup of PI, and (b) the increased latency associated with the buildup of PI reflects a slowing of the exponential retrieval process and not a delay in the onset of recall. With regard to the first of these results, a natural question to consider is why these retrieval curves exhibit the form they do. More specifically, what does this finding imply about the nature of retrieval that would not be true if the curves were linear or hyperbolic? The property of a system that gives rise to an exponential function has been described in various ways, the most common being (somewhat ironically) *memoryless* (e.g., Cinclar, 1975; Papoulis, 1984). This term is intended to emphasize the fact that the momentary probability of retrieval associated with individual items remains constant throughout the recall period and is not affected by preceding events. Thus, exponential retrieval is not consistent with the idea of a linear scan of a memory set (such as that often assumed to operate in the Sternberg task) or with the idea that items are becoming less likely to be recalled because of the previous retrieval of other items (i.e., output interference). With regard to the former, the momentary probability of recall associated with the not-yet-retrieved items increases as other items are recalled. With regard to the latter, the momentary probability of recall associated with those items decreases as other items are recalled. In neither case would the resulting retrieval curve be exponential in form.

Researchers have realized for many years that the simplest scheme consistent with exponential retrieval involves sampling with replacement from a memory search set (Bousfield & Sedgewick, 1944). The search set itself consists of some or all of the items from the list and, perhaps, some number of additional items that were not on the list. If a constant rate of sampling is assumed, the average time to recall simply reflects the size of that search set (cf. McGill, 1963). Thus, interpreted in light of this model, the present results are consistent with a central assumption of temporal discrimination theory, namely, that subjects are searching through an enlarged search set with each passing trial. The Trial 4 data further suggest that the size of the search set narrows with the provision of a before or an after cue.

The latter point regarding the effect of the after cue is significant because it suggests that the role of the after cue may differ from what was previously thought. Gardiner et al. (1972) suggested that the after cue might function to solve the temporal discrimination problem by uniquely identifying which of the many items under consideration appeared on Trial 4. If so, one would still expect to see a long retrieval latency because the search would involve items from all four trials. Instead, the present results suggest that the before and after cues may operate in the same way, namely, by narrowing the search to the appropriate set of items.

The absence of an effect of PI on the onset of recall (as reflected by the constant value of c) rules out a variety of alternative search scenarios. For example, one possible explanation for the increased recall latency associated with the buildup of PI is that the delay reflects the time taken to find the appropriate area to search and not the time taken to retrieve the individual items once that area is found. That search onset might be affected under some conditions does not seem altogether unreasonable. For example, the delay in recovering memories associated with a meal eaten 2 weeks ago relative to one eaten today may reflect the extra time required to find the appropriate area to search. Once that is done, however, the retrieval of individual memories (e.g., who was there, what topics were discussed, what was eaten) may proceed at a relatively quick pace. On the basis of the present results, however, PI does not affect the time taken to find the appropriate search set.

A second search scenario that would mistakenly predict an increase in c without a corresponding increase in τ holds that, during the buildup of PI, items from every trial are searched in sequential fashion. Thus, for example, on Trial 2, subjects might first search the items from Trial 1 (because those were the first to be associated with the retrieval cue) followed by a search of the items from Trial 2. Once again, the obtained pattern of results does not support this idea. Instead, the results are more consistent with the idea that the search set expands with each trial until a new category is introduced.

Experiment 2

The previous experiment suggested that the buildup of PI results in a retardation of an exponential retrieval process, which can then be accelerated by the provision of appropriate before and after cues. A question that remains unanswered is whether the source of this effect occurs at the time the items are encoded or during the course of the retention interval. The temporal discrimination theory of PI appears to predict that the search set will increase with each passing trial regardless of the size of the retention interval. Thus, for example, on Trial 2 the subject must perform the relatively easy discrimination of deciding which items appeared on Trial 1 and which appeared on Trial 2. Nevertheless, the set of items under consideration on Trial 2 is still larger than that on Trial 1. In the second experiment we investigated this issue by measuring the time course of retrieval during the buildup of PI after short and long retention intervals.

Method

Subjects. The subjects were 42 undergraduates of the University of California, San Diego. Their participation in the experiment satisfied a psychology course requirement.

Materials and design. The materials used in this experiment were the same as those used in Experiment 1 (Table 2). However, unlike in the previous experiment, subjects were presented with six sets of three trials, and no subtle category shifts occurred. Each set of three trials involved words from one of the six subcategories presented in Table 2. The subcategories were presented in one of six random orders. For Sets 1 through 3, unrelated category subsets were used (e.g., indoor sports, coastal states, and paired body parts). For Sets 4 through 6, the complementary subsets were presented in the same order (e.g., outdoor sports, inland states, and singular body parts). In all cases, a before cue was presented that specified the subcategory to which the upcoming words belonged.

Two retention intervals (3 s or 27 s) were programmed according to one of six random orders. The distribution of retention intervals was such that half of the lists were followed by a short retention interval and half were followed by a long retention interval.

Procedure. The procedure was otherwise identical to that of Experiment 1 except that before cues were always provided and after cues were never provided. Thus, unlike in Experiment 1, the recall period began immediately following the retention interval. In all, subjects completed 6 practice trials and 18 experimental trials. During the 20-s recall periods, recall responses were again timed. A slightly different timing strategy was implemented to improve accuracy and efficiency of data collection. Specifically, as each item was recalled, an experimenter tapped a key on the keyboard, and the computer timed the response.

Results and Discussion

Figure 4 shows the proportion of items recalled in each condition as a function of trial. These data reflect the typical pattern of results: Performance remains constant at the short retention interval and declines precipitously at the long retention interval (cf. Keppel & Underwood, 1962). An ANOVA performed on these data revealed significant main effects for retention interval and trial, $F(1, 41) = 117.75$, $MS_e = 1.35$, and $F(2, 82) = 10.84$, $MS_e = 1.79$, respectively, as well as a significant interaction, $F(2, 82) = 11.11$, $MS_e = 1.46$.

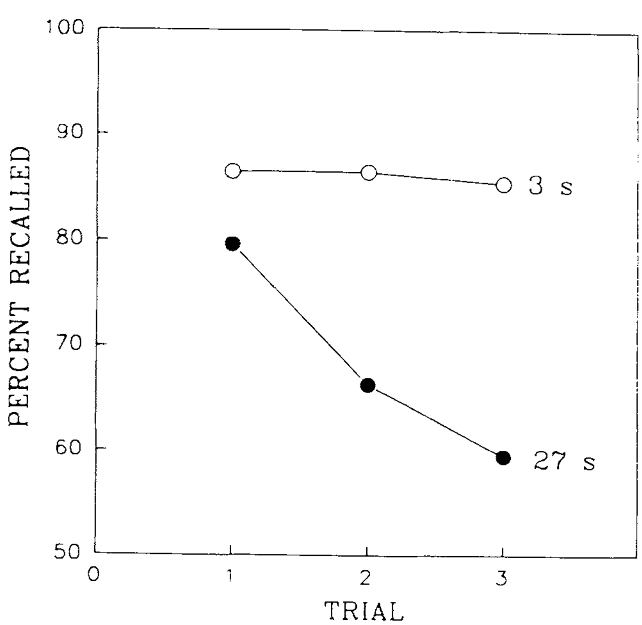

Figure 4. Percentage of items recalled after the short (3-s) and long (27-s) retention intervals as a function of trial.

The question of interest is whether the effect of PI on the recall latency distribution differs following short and long retention intervals. In Experiment 1, subjects apparently initiated recall just prior to the nominal onset of the recall period (that is, during the period of time reserved for the presentation of the after cue). In Experiment 2, the recall period began immediately following the retention interval, which allowed for the more accurate timing of the fastest recall responses. This method revealed a generally overlooked property of the dynamics of free recall. More specifically, the latency distributions were characterized by a sharply ascending arm followed by the expected exponential decay. We interpret this result as arising from variability in the onset of exponential retrieval. That is, on most trials, subjects began recalling in the second 1-s bin, but on occasional trials they began recalling in the first or third bin instead.

If one assumes that recall RT involves one or more normally distributed stages (e.g., finding a search set) followed by an exponential stage (e.g., retrieval from that set), then the recall latency distribution should be described by the convolution of the normal and exponential distributions (Hohle, 1965). This distribution, known as the ex-Gaussian, has previously been shown to accurately describe recognition latency distributions (e.g., Heathcote et al., 1991; Hockley, 1984; Luce, 1986). The somewhat forbidding mathematical form of this conceptually simple function is

$$f(t) = \frac{e^{\frac{1}{2}(\sigma/\tau)^2 - (t - \mu)/\tau}}{\tau\sqrt{2\pi}} \int_{-\infty}^{(t - \mu)/\sigma - (\sigma/\tau)} e^{-y^2/2}\, dy$$

where τ represents the average time of the exponential stage, and μ and σ represent the mean and standard deviation of the normally distributed stages, respectively. Mean recall latency is equal to the sum of μ and τ.

Figure 5 shows the noncumulative distributions from this experiment along with the best-fitting ex-Gaussian equation. The three graphs in the left panel display the data from Trials 1–3 (top to bottom) at the 3-s retention interval, whereas the three graphs in the right panel display the corresponding data for the 27-s retention interval. The figure also presents the obtained chi-square values for each fit along with the corresponding degrees of freedom. Although a significant chi-square was observed in one case (Trial 2 of the 3-s condition), the ex-Gaussian distribution appears to provide an accurate description of these data. By way of comparison, alternative distributions fitted to the data (specifically the gamma and Weibull distributions) produced significant chi-square values in three or more cases each. However, the lognormal distribution rivaled the ex-Gaussian (producing only one significant chi-square), which is not surprising given that both the ex-Gaussian and lognormal distributions fit recognition latency distributions reasonably well (Ratcliff & Murdock, 1976).

The parameter estimates for each of the curves plotted in Figure 5 are shown in Table 5. These data reveal that the mean and standard deviation associated with the Gaussian process (μ and σ, respectively) remained constant across trials at the short retention interval and increased slightly (although not significantly) at the long retention interval.

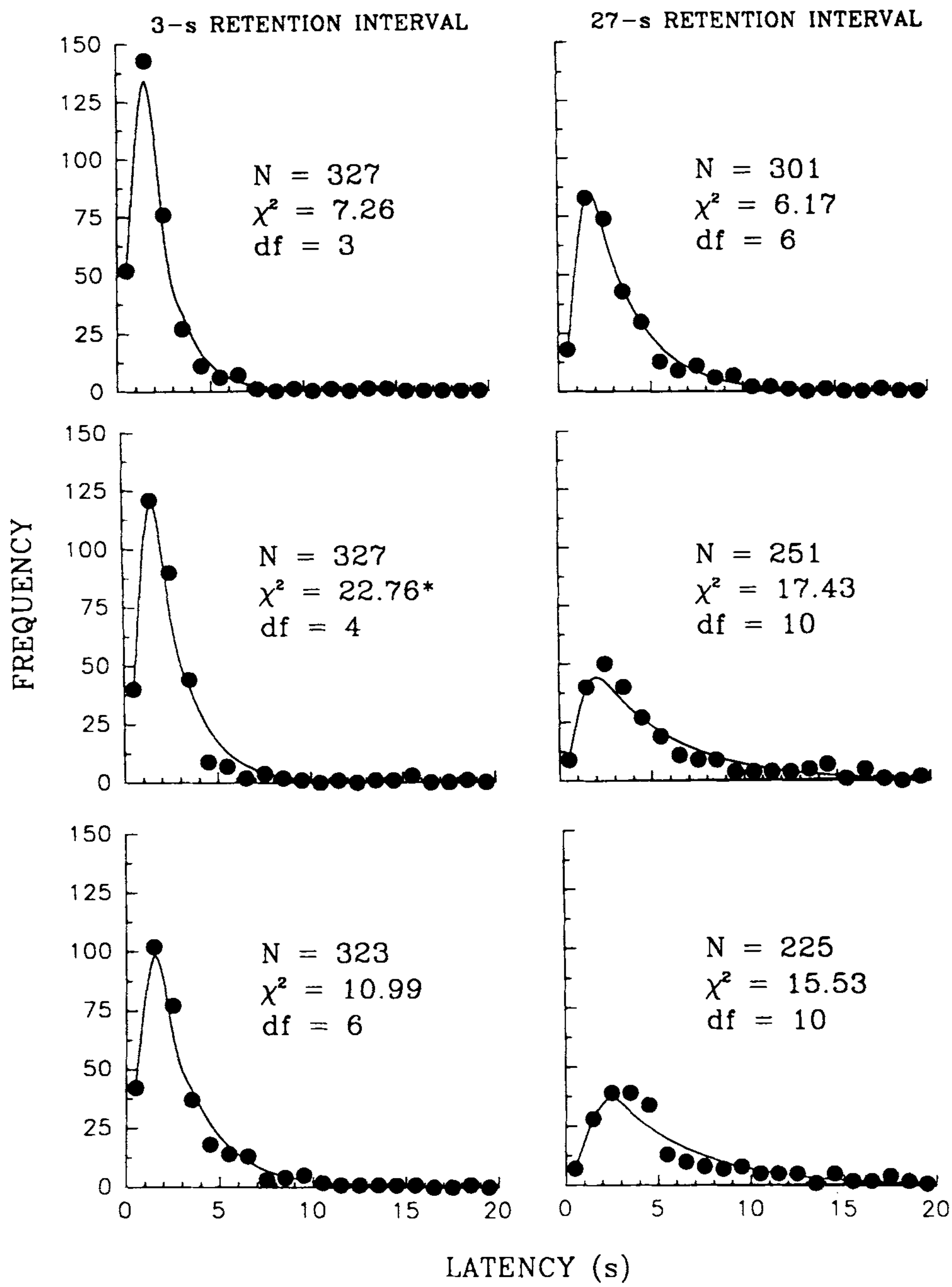

Figure 5. Recall latency distributions for Trials 1, 2, and 3 (top to bottom) after the 3-s retention interval (left panel) and recall latency distributions for Trials 1, 2, and 3 (top to bottom) following the 27-s retention interval (right panel). (The solid curves represent the best-fitting ex-Gaussian distribution. The asterisk indicates a statistically significant deviation from the ex-Gaussian, $\alpha = .05$.)

Note that this finding replicates an important conclusion from Experiment 1. Had the buildup of PI affected recall onset latency, this should have been reflected in an increase in μ (and, of course, would be reflected by an increase in c were the distributions plotted cumulatively).

In agreement with the results of Experiment 1, the average latency associated with the exponential process (τ) increased with each passing trial at the 27-s retention interval. Of more interest, however, are the results from the 3-s retention interval. Although the number of items recalled remained essentially constant (see Figure 4), the average exponential latency (τ) increased with each passing trial. The 1.42-s average latency on Trial 1 was exceeded by the 1.76-s latency obtained on Trial 2 (though the difference was not signifi-

Table 5
Ex-Gaussian Parameter Estimates for the Retrieval Functions From Experiment 2

Trial	Retention interval (seconds)	τ	μ	σ
1	3	1.42 (0.9)	0.72 (.07)	0.34 (.11)
2	3	1.76 (.13)	0.79 (.11)	0.40 (.16)
3	3	2.27 (.14)	0.66 (.10)	0.36 (.22)
1	27	2.42 (.18)	0.98 (.11)	0.48 (.10)
2	27	4.21 (.35)	1.06 (.14)	0.56 (.13)
3	27	4.40 (.40)	1.25 (.17)	0.70 (.17)

Note. The parenthetical values represent the asymptotic standard errors of the parameter estimates. τ = average time of the exponential stage. μ = mean of the normally distributed stages. σ = the standard deviation of the normally distributed stages.

cant), which, in turn, was significantly exceeded by the 2.27-s latency obtained on Trial 3, $t(10) = 2.65$. As might be expected, the difference between Trials 1 and 3 was highly significant, $t(9) = 4.79$.

The latter finding is important for several reasons. First, in their analysis of recall latency and probability measures, MacLeod and Nelson (1984) reported that they had not found a single valid instance in which recall latency was more sensitive to an experimental manipulation than was probability of recall. The only apparent exceptions were rendered uninterpretable by the presence of a ceiling effect for the recall probability data. As indicated above, the 3-s retention interval data reveal the elusive pattern described by MacLeod and Nelson: Probability of recall remained constant while latency increased. Thus, in this case, latency to recall appears to provide a more sensitive measure than probability.

Could this simply reflect another case of a recall ceiling effect? Although that possibility cannot be ruled out entirely, it seems unlikely for three reasons. First, performance never exceeded 86.5%, which is below the low end of the 90%–100% range often used as a rule-of-thumb ceiling. Second, we used a fast rate of presentation in these experiments (two words per second) precisely because a slower rate led to an actual ceiling effect on Trial 1 (with performance in the vicinity of 95% correct). Third, although some reduction in variability is apparent, the data do not exhibit the dramatic constriction one would expect if performance were constrained against an artificial ceiling. For Trials 1–3 at the 3-s retention interval, the mean numbers of items recalled (out of 9 across the three trials) were 7.79, 7.79, and 7.69 with corresponding standard deviations of 1.30, 1.03, and 1.30. For Trials 1–3 at the 27-s retention interval, the mean numbers of items recalled were 7.17, 5.98, and 5.36 with corresponding standard deviations of 1.10, 1.69, and 1.68. The Trial 1 data are the most telling because the obtained variability in the 3-s condition (with performance at 86.5% correct) actually exceeded that in the 27-s condition (with performance at 79.7% correct). Such a finding would be unlikely if the 3-s data were compressed against a ceiling. Nevertheless, the larger variability for Trials 2 and 3 of the 27-s condition suggest that some constriction may exist (although this constriction is evident even when performance is as low as 79.7% correct).

If we assume the absence of a ceiling effect, then the present results suggest that latency measures can be more sensitive to PI than probability measures. This should not be too surprising if it is true that these measures reflect different properties of retrieval (the central conclusion reached by MacLeod & Nelson, 1984). As shown by Keppel and Underwood (1962), PI does not affect recall probability at a short delay. However, if, as suggested by temporal discrimination theory, subjects search through a larger set of items on later trials, then latency to recall should still increase (i.e., in this case, latency should be more sensitive than probability).

Certain aspects of the present results bear on a temporal distinctiveness theory of retrieval advanced by Glenberg and Swanson (1986). That theory assumes that temporal cues are used to construct a search set and that the size of the retention interval determines how large that search set will be. When the retention interval is short, high-frequency components of the temporally changing context are more or less uniquely associated with the list items. Because those components are still present following a short retention interval, they can be used to construct a search set that consists mainly of the relevant list items. After a long retention interval, by contrast, those high-frequency components have long since passed. Thus, more stable aspects of the temporal context that were present during list presentation (and which are still present after a long retention interval) are instead used to establish a search set. Because those contextual features are associated with the list items as well as many other events, the size of the search set will be correspondingly larger. This theory may explain why release from PI can occur by merely increasing the interval of time separating Trial 3 from Trial 4 (Kincaid & Wickens, 1970; Loess & Waugh, 1967).

If Glenberg and Swanson's (1986) temporal distinctiveness theory is correct, then one would expect to find an increase in the exponential parameter, τ, as the retention interval increases (reflecting a larger search set). As shown in Table 5, the retention interval manipulation resulted in an increase in both μ and τ for all three trials (cf. Dosher, 1981). The effect on τ was significant in all three cases, $t(9) = 4.54$, $t(14) = 5.70$, and $t(16) = 4.60$ for Trials 1–3 respectively, but the effect on μ reached significance only for Trial 3, $t(16) = 2.83$. Whereas the increase in τ is consistent with a growing search set, the increase in μ suggests that, after a long retention interval, subjects might require more time to access the appropriate area to search.

Two methodological issues that might be raised at this point concern the extent to which measurement error may have played a role in affecting the shape of the obtained distributions and the extent to which those distributions are representative of individual subjects. To address these issues, we obtained response latency distributions from 4 subjects who were each exposed to a sufficient number of recall trials to obtain smooth individual distributions. For these subjects, trials consisted of the presentation of five-item lists of high-frequency one-syllable words (presented at a rate of one word every 2 s), an 18-s retention interval filled with a distractor task, and a 20-s recall period. The subjects participated in two sessions each, and they received 30 lists per session. Recall responses were timed as before (i.e., the experimenter tapped a computer key) as well as by a voice-activated relay attached to a millisecond clock. Subjects were instructed to enunciate each word without slurring them together (which would cause them to be missed by the voice key). With minimal practice, subjects were able to comply with this instruction without difficulty.

One drawback to using a voice key is that it fails to distinguish between correct recalls and intrusions. However, the experimenter did keep track of intrusions (by tapping a different key for errors), which allowed the distributions to be plotted with or without the intrusions. Intrusions, which typically occurred late in the recall period, were extremely rare (32 out of 865 responses), and the results did not differ in either case.

The individual latency distributions obtained from the 4 subjects (using the voice-activated relay) are presented in Figure 6. All four plots exhibit the general ex-Gaussian form, although the deviations were significant in one case. The estimated values of τ were very similar whether the latencies were measured by the voice key or by the experimenter. The small differences that were observed did not approach significance in any case. As might be expected, however, μ was slightly longer for the experimenter-timed data. For the aggregate fit using the data from all 4 subjects, the estimated values of τ were 2.92 s and 2.93 s for

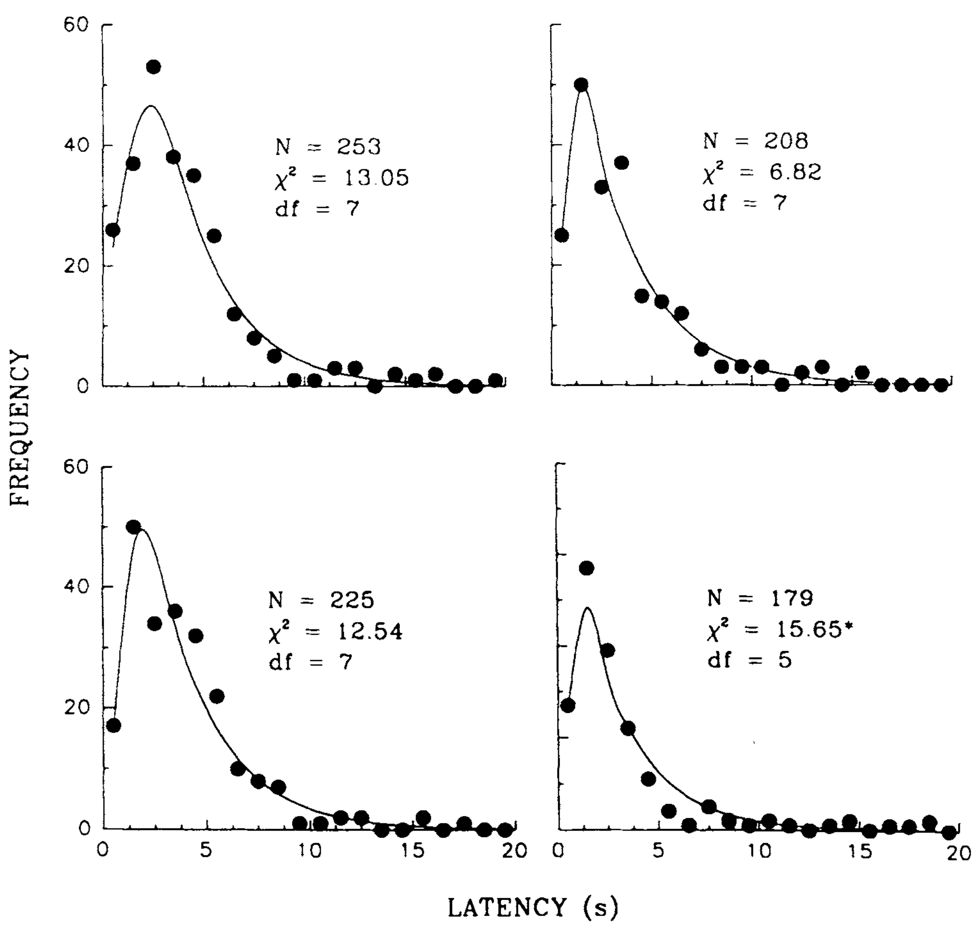

Figure 6. Recall latency distributions from 4 individual subjects. (The solid curves represent the best-fitting ex-Gaussian distribution. The asterisk indicates a statistically significant deviation from the ex-Gaussian, α = .05.)

the voice key and experimenter-timed data, respectively, whereas the corresponding values of μ were 0.74 s and 1.18 s (the 440-ms difference presumably reflecting experimenter RT). Thus, although experimenter-timed distributions are right-shifted by a small amount, they nevertheless appear to provide an otherwise undistorted picture of the dynamics of free recall.

General Discussion

Previous analyses of free-recall latency distributions are extremely scarce, and none have involved the kind of parametric analysis offered here. Thus, the present results provide the first demonstration that (a) recall latency distributions, like recognition distributions, are accurately described by the ex-Gaussian; (b) the buildup of PI on the Brown–Peterson task is associated with an increase in the exponential parameter of the ex-Gaussian, τ, without a corresponding increase in the mean of the Gaussian component, μ; and (c) increasing the retention interval from 3 to 27 s results in an increase in both τ and μ (though the latter result must be regarded as tentative). The importance of a distributional analysis of recognition latency has been emphasized for years, and a recent review suggested that analyses that are based on summary statistics are not only incomplete but also potentially misleading (Heathcote et al., 1991). Exactly the same arguments apply to the analysis of recall latency distributions because mean latency, which is equal to the sum of μ and τ, does not reveal whether one or both of these parameters changed as a result of some experimental manipulation (or even whether the two parameters changed in the same direction).

Theoretical Interpretation

A theoretical analysis of the present results should be able to explain both the mathematical form of the obtained distribution as well as the way in which the parameters change under the influence of PI. One general model consistent with an ex-Gaussian distribution holds that the contents of a search set on a given trial are established by a retrieval cue (such as the category name) and that the time required to complete that process affects μ only. After that, items are retrieved from the search set according to an exponential process. The class of models consistent with an exponential retrieval stage share the assumption that the momentary probability of retrieving individual items from a search set remains constant. Thus, for example, a retrieval cue might simultaneously activate a set of representations (including some subset of items from the list plus some extralist items) such that each has the same momentary probability, p, of reaching conscious awareness. If the value of p remains more or less constant during the recall period, then retrieval would follow an exponential time course.

One might further assume that the more items activated by the retrieval cue, the lower the momentary probability of retrieval associated with any individual item (which would be reflected by a larger τ). For example, in a sampling-with-replacement serial search model, the average time required to find target items in a search set increases linearly with the size of that set (McGill, 1963). The analogous parallel search model would hold that the momentary probability of summoning an item to conscious awareness decreases (and recall latency increases) as the cue's retrieval strength is distributed over a larger number of items. In either case, the increase in τ associated with the buildup of PI can be reasonably interpreted as reflecting an increase in the size of that search set (a central assumption of temporal discrimination theory). Because the search set on Trial 1 consists of items from that trial only, the average time required to sample all of them at least once (viz., τ) is relatively short. On Trial 2, the search set expands to include items from both Trial 1 and Trial 2 and thus requires more time to search (i.e., τ increases). On Trial 3, the search set increases further, causing still greater increases in recall latency.

Although this analysis provides one account of the growth in τ associated with the buildup of PI, it does not explain the corresponding decrease in the absolute probability of recall. Indeed, an explanation for this effect is somewhat elusive. Temporal discrimination theory holds that performance declines on later trials because of the difficulty involved in mentally separating items that appeared on the most recent trial from those that appeared on earlier trials (and not because the most recent items are difficult to retrieve from memory). This widely accepted theory provides a particularly compelling explanation for the results shown in Figure 4. When the retention interval is short, the temporal discrimination is relatively easy and performance on later trials does not suffer. As the retention interval increases, however, that discrimination becomes increasingly difficult and performance suffers accordingly.

In spite of the ability of temporal discrimination theory to explain some important PI phenomena, Dillon and his associates concluded that the reduction in recall probability on later trials occurs because of search failure, not because of temporal discrimination difficulties (Dillon, 1973; Dillon & Bittner, 1975; Dillon & Thomas, 1975). They arrived at this conclusion because a variety of interventions designed to solve the temporal discrimination problem (e.g., allowing subjects to recall all of the items that came to mind) failed to have a significant effect on performance. Furthermore, when confidence ratings were taken, subjects tended to exhibit high confidence in their responses from the current trial and low confidence in intrusions from earlier lists (which suggests a relatively accurate temporal discrimination).

The present results do not resolve this issue, but they do suggest that, in either case, a larger area of memory may be searched as PI builds. Many earlier investigations of PI have shown that items from previous trials are still available on later trials. Indeed, Dillon himself found that when subjects were asked to recall all of the items that came to mind on each trial, their responses included items from the current trial as well as many from the preceding trials (Dillon & Thomas, 1975). The present results suggest not only that those items are available but that they may be searched in a more or less obligatory way on each trial and thereby result in an increased recall latency.

Comments on the Exponential Form of Retrieval

For several reasons, the exponential form of retrieval should be somewhat surprising. As mentioned earlier, one implication of this finding is that the momentary probability of retrieval associated with individual items remains constant over the course of the recall period. At the very least, one might expect the process of forgetting to cause items to become less accessible with each passing moment. However, verbal forgetting functions flatten out very quickly (cf. Wixted & Ebbesen, 1991). Thus, to a reasonable approximation, the forgetting that occurs over the course of a brief recall period may be negligible.

Beyond the process of forgetting, output interference also seems to require that as each item is retrieved, the probability of retrieving the remaining items declines. Indeed, Rundus (1973) explicitly incorporated this assumption into his memory search model. Furthermore, Bousfield et al. (1954) appealed to output interference to explain why the retrieval curves obtained in that experiment were better described by the hyperbola than the exponential. Unlike exponential recall, hyperbolic recall implies a diminishing momentary probability of retrieval and is thus consistent with the presence of output interference. However, the hyperbola did not describe the retrieval curves obtained in the present experiments, which seems to suggest that output interference did not play a major role.

How can the idea of constant retrieval probability be reconciled with the abundant evidence for output interference? There are at least two possibilities. First, it should be clear that the most compelling evidence for output interference from secondary memory does not suggest that as each item is recalled, the probability of retrieving the remaining ones decreases. Instead, most of the evidence suggests that as items from one category of a multicategory list are recalled, the probability of subsequently retrieving items from another category declines. For example, Smith (1971, 1973), and more recently Roediger and Schmidt (1980), found that the number of items recalled per category declined with the category's output position even when the effect of retention interval was taken into account. This result does not necessarily imply that, within a category, the retrieval of one item negatively affects the retrieval of others. Instead, recalling items from the first category may alter the search set established by the second category such that it embraces fewer list items (cf. Sloman, Bower, & Rohrer, 1991). If that were the case, then the probability of recall would be reduced relative to the first category (i.e., output interference would be observed), but the time course of retrieval would be exponential in both cases. A similar explanation might apply to the effects of part-list cues (e.g., Slamecka, 1968, 1969), namely, that such cues affect item accessibility by altering the boundaries of the subsequent search without affecting the momentary probability of retrieval associated with the items contained within those boundaries.

Alternatively, if one instead accepts the idea that the recall of each item from a search set affects the subsequent probability of recall associated with other items in the same search set, then output interference and exponential recall can be reconciled in another way. If the successful retrieval of one item creates a new copy of that item in the search set (or a new link between that item and the retrieval cue) that essentially replaces another item in the search set, then the evicted items would not be retrieved more slowly (thereby changing the form of the exponential retrieval function) because they would not be retrieved at all. Such a retrieval scheme could produce an exactly exponential function, depending on specific assumptions about the contents of the original search set.

Although the exponential provided a reasonably accurate description of the data presented here, it should be clear that the same result would not be expected under all circumstances. Indeed, any manipulation that introduces widely differing retrieval probabilities among the list items should produce a function other than an exponential. Order effects (e.g., recalling primacy items first) and item clustering (e.g., recalling *bread* and *butter* together) are two such examples. These effects can be accommodated by the models discussed earlier if one assumes that sets of related items are retrieved as a functional unit (e.g., a single draw retrieves both *bread* and *butter*) and if one assumes that items receiving the most rehearsal are more likely to be multiply represented (e.g., Bernbach, 1970; Laming, 1992). To the extent that either of these effects are pronounced, however, nonexponential retrieval would be expected.

The present experiments involved short lists of rapidly presented items followed by a retention interval filled with a relatively demanding distractor task. Such conditions probably minimize both clustering and order effects. However, both effects become much more pronounced in multitrial free recall involving long lists of words (Bousfield, Puff, & Cowan, 1964; Tulving, 1962). That is, when the same list is presented repeatedly for recall, the subject's output becomes increasingly systematic and ordered. Although the issue has never been investigated, one might expect to find that the resulting recall latency data will be less accurately described by the ex-Gaussian distribution on later trials as clustering and order effects begin to dominate.

In a recent article, Laming (1992) found that recall latency data reported by Peterson and Peterson (1959) were not accurately described by the exponential in two of six conditions. Thus, in some cases at least, latency data may not conform to an exponential model (indeed, one condition of our Experiment 2 yielded a significant chi-square as well). As an alternative to an exponential model, Laming presented a fairly comprehensive theory of performance on the Brown–Peterson task, which assumes that the presentation of a single list item creates a memory trace that subsequently becomes less accessible according to an inverse function of time. Any covert rehearsals that happen to occur during the subsequent distractor task create additional traces that become less accessible in the same way. This model offered a reasonably accurate description of the Peterson and Peterson recall latency data, but, because the model has not been extended to the situation studied here (viz., multiitem lists in a PI paradigm), its ability to describe the present distributions is not clear. Nevertheless, the eventual comparison of that model with the ex-Gaussian distribution represents an obvious next step.

The Role of Decision Processes

We have to this point ignored one possible contributor to the shape of the obtained density functions: the role of decision processes. Indeed, generate–recognize theories explicitly assume that decision processes play an important role in free recall (e.g., Anderson & Bower, 1972; Kintsch, 1970). According to this account, candidate items are first retrieved from memory and then are individually evaluated for possible overt recall. Conceivably, the latency increases observed in the present experiments could have resulted from an increase in the decision component rather than from an increase in the search component.

Although this idea cannot be ruled out entirely, it seems unlikely in view of evidence suggesting that large variations in decision thresholds have virtually no effect on free recall. Roediger and Payne (1985) observed that a stringent decision threshold is often assumed to operate in free recall because subjects typically make few overt intrusions. To test this idea, they varied the decision threshold by giving different recall instructions to different groups of subjects. The first group was given ordinary free-recall instructions, a second group was instructed to recall any words that came to mind, and a third group was instructed to write down a total of 50 words even if they had to guess. The probability of recall was unaffected by these instructions (although the latter groups generated many more intrusions). As indicated earlier, a similar result was reported by Dillon and Thomas (1975) in the context of a PI experiment. That is, lowering the decision threshold did not improve recall on later trials but did greatly increase the number of intrusions from previous trials. On the basis of such findings, Roediger and Payne concluded that subjects apparently do not generate items and then subject them to a recognition decision. Instead, it seems that the items that are retrieved into conscious awareness are overtly recalled without much deliberation.

The possibility nevertheless remains that variations in the decision threshold would affect recall latency without affecting recall probability. This issue could be tested by varying the decision threshold and carefully timing the recall latencies. We would expect to find no effect on the basis of the previous evidence suggesting that free recall does not entail a significant decision component. Nevertheless, the question cannot be completely resolved until the appropriate experiments are performed.

Recall and Recognition Latency Distributions

The ex-Gaussian distribution, which has now been shown to provide a reasonably accurate description of recall latency, has previously been shown to accurately describe recognition latency as well. This might be a purely coincidental result, or it might reflect an underlying commonality. Indeed, in light of numerous theories suggesting that recognition involves a retrieval component, the connection between these two latency phenomena might not be surprising. For example, Mandler's (1980) dual-process theory of recognition holds that slower recognition decisions (i.e., those comprising the exponential tail of the ex-Gaussian) involve retrieval, whereas the faster ones are based solely on familiarity (e.g., Mandler & Boeck, 1974). If that account is true, then one might expect to find that the tail of recall and recognition latency distributions would be described by the same function (viz., the exponential). More important, the τ parameter of the exponential should be affected in similar ways whether one examines recall or recognition latency distributions. As yet, a PI analysis of the recognition latency distribution has not been performed. If the exponential tail of the recognition latency distribution reflects retrieval, and if the search set increases with the buildup of PI, then τ should increase with each passing trial.

The analysis of free-recall latency distributions represents a largely unexplored region of human memory research. The present findings suggest that the further empirical analysis of these distributions may be worthwhile, especially as new theories of human memory (e.g., Laming, 1992) reflect an increased interest in the subject. Shiffrin (1970) long ago observed that "latencies provide a powerful tool for the examination of the characteristics of the search through the long-term store" (p. 440). These words are as true today as they were then.

References

Anderson, J. R. (1981). Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 326–343.

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

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. R. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Learning, memory, and thinking* (Vol. 1, pp. 243–293). New York: Freeman.

Baddeley, A. (1990). *Human memory: Theory and practice.* Needham Heights, MA: Allyn & Bacon.

Bernbach, H. A. (1970). A multiple-copy model for postperceptual memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 103–116). San Diego, CA: Academic Press.

Bousfield, W. A., & Barclay, W. D. (1950). The relationship between order and frequency of occurrence of restricted associated responses. *Journal of Experimental Psychology, 40,* 643–647.

Bousfield, W. A., Cohen, B. H., & Silva, J. G. (1956). The extension of Marbe's law to the recall of stimulus words. *American Journal of Psychology, 69,* 429–433.

Bousfield, W. A., Puff, C. R., & Cowan, T. M. (1964). The development of constancies in sequential organization during repeated free recall. *Journal of Verbal Learning and Verbal Behavior, 3,* 489–495.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of restricted associative responses. *The Journal of General Psychology, 30,* 149–165.

Bousfield, W. A., Sedgewick, C. H. W., & Cohen, B. H. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology, 67,* 111–118.

Bousfield, W. A., Whitmarsh, G. A., & Esterson, J. (1958). Serial position effects and the "Marbe Effect" in the free recall of mean-

ingful words. *The Journal of General Psychology, 59,* 255–262.
Cinclar, E. (1975). *Introduction to stochastic processes.* Englewood Cliffs, NJ: Prentice Hall.
Crowder, R. G. (1976). *Principles of learning and memory.* New York: Wiley.
Dillon, R. F. (1973). Locus of proactive interference effects in short-term memory. *Journal of Experimental Psychology, 99,* 75–81.
Dillon, R. F., & Bittner, L. A. (1975). Analysis of retrieval cues and release from PI. *Journal of Verbal Learning and Verbal Behavior, 14,* 616–622.
Dillon, R. F., & Thomas, H. (1975). The role of response confusion in proactive interference. *Journal of Verbal Learning and Verbal Behavior, 14,* 603–615.
Dosher, B. A. (1981). The effects of delay and interference: A speed–accuracy study. *Cognitive Psychology, 13,* 551–582.
Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior, 11,* 778–783.
Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 3–15.
Gorfein, D. S., & Jacobson, D. E. (1973). Memory search in a Brown–Peterson short-term memory paradigm. *Journal of Experimental Psychology, 99,* 82–87.
Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347.
Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6,* 598–615.
Hohle, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology, 69,* 382–386.
Indow, T., & Togano, K. (1970). On retrieving sequence from long-term memory. *Psychological Review, 77,* 317–331.
Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153–161.
Kincaid, J. P., & Wickens, D. D. (1970). Temporal gradient of release from proactive inhibition. *Journal of Experimental Psychology, 86,* 313–316.
Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory* (pp. 333–373). San Diego, CA: Academic Press.
Laming, D. (1992). The analysis of short-term retention: Models for Brown–Peterson experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1342–1365.
Loess, H., & Waugh, N. C. (1967). Short-term memory and intertrial interval. *Journal of Verbal Learning and Verbal Behavior, 6,* 455–460.
Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.
MacLeod, C. M., & Nelson, T. O. (1984). Response latency and response accuracy as measures of memory. *Acta Psychologica, 57,* 215–235.
Maindonald, J. H. (1984). *Statistical computation.* New York: Wiley.
Mandler, G. (1980). Recognizing: The judgment of prior occurrence. *Psychological Review, 87,* 252–271.
Mandler, G., & Boeck, W. J. (1974). Retrieval processes in recognition. *Memory & Cognition, 2,* 613–615.
McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309–360). New York: Wiley.
O'Neill, M. E., Sutcliffe, J. A., & Tulving, E. (1976). Retrieval cues and release from proactive inhibition. *American Journal of Psychology, 89,* 535–543.
Papoulis, A. (1984). *Probability, random variables, and stochastic processes* (2nd ed.). New York: McGraw-Hill.
Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 58,* 193–198.
Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review, 83,* 190–214.
Ratkowsky, D. A. (1983). *Nonlinear regression modeling: A unified practical approach.* New York: Marcel Dekker.
Roediger, H. L., III, & Payne, D. G. (1985). Recall criterion does not affect recall level or hypermnesia: A puzzle for generate/recognize theories. *Memory & Cognition, 13,* 1–7.
Roediger, H. L., III, & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired-associate lists. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 91–105.
Roediger, H. L., III, Stellon, C. C., & Tulving, E. (1977). Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 174–188.
Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12,* 43–50.
Shiffrin, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of human memory* (pp. 375–447). San Diego, CA: Academic Press.
Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76,* 504–513.
Slamecka, N. J. (1969). Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology, 81,* 557–560.
Sloman, S. A., Bower, G. H., & Rohrer, D. (1991). Congruency effects in part-list cuing inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 974–982.
Smith, A. D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior, 10,* 400–408.
Smith, A. D. (1973). Input order and output interference in organized recall. *Journal of Experimental Psychology, 100,* 147–150.
Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University Press.
Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69,* 344–354.
Underwood, B. J. (1945). The effect of successive interpolations on retroactive and proactive inhibition. *Psychological Monographs, 59* (Whole No. 273).
Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49–60.
Wickens, D. D. (1972). Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 195–215). Washington, DC: Winston.
Wilkes, S. S. (1962). *Mathematical statistics.* New York: Wiley.
Wixted, J. T., & Ebbesen, E. (1991). On the form of forgetting. *Psychological Science, 2,* 409–415.

Received February 21, 1992
Revision received December 31, 1992
Accepted January 14, 1993 ■