An analysis of latency and interresponse time in free recall

DOUG ROHRER and JOHN T. WIXTED
University of California, San Diego, La Jolla, California

In four experiments, subjects freely recalled previously studied items while a voice key and computer recorded each item's recall latency relative to the onset of the recall period. The measures of recall probability and mean recall latency were shown to be empirically independent, demonstrating that there exists no a priori relationship between the two. In all four experiments, latency distributions were fit well by the ex-Gaussian, suggesting that retrieval includes a brief normally distributed initiation stage followed by a longer exponentially distributed search stage. Further, the variation in mean latency stemmed from the variation in the duration of the search stage, not the initiation stage. Interresponse times (IRTs), the time elapsed between two successive item recalls, were analyzed as well. The growth of mean IRTs, plotted as a function of output position, was shown to be a simple function of the number of items not yet recalled. Finally, the mathematical nature of both free recall latency and IRT growth are shown to be consistent with a simple theoretical account of retrieval that depicts mean recall latency as a measure of the breadth of search.

Perhaps the most striking feature of episodic free recall is the distinctively long pause that sometimes precedes a response. Even a study list with only a few items often results in response latencies greater than 10 sec. These lengthy response times are unique to free recall, as cued recall and recognition typically yield response times of 1 or 2 sec. Yet although analyses of cued recall latency (e.g., MacLeod & Nelson, 1984) or recognition latency (e.g., Hockley, 1982) are quite common, few investigators have measured free recall latency, the time elapsed from the onset of the recall period to the recall of an item, or interresponse times (IRTs), the time between consecutive retrievals.

The limited attention given to either temporal measure may stem from the belief that recall probability and recall latency are a priori inversely related and therefore redundant measures. That is, an experimental manipulation that weakens the memory trace and thereby reduces recall probability should necessarily increase recall latency as well. If this were the case, researchers could simply ignore free recall latency without overlooking any important information. Although such negative correlations can occur, there is no prior evidence that they either must (or even usually) occur. Indeed, three experiments in the present investigation yielded correlations between recall probability and latency that were negative, nonexistent,

and positive, thus illustrating that the measure of free recall latency provides unique information.

While the measure of recall latency is important in its own right, the empirical form of recall latencies also suggests a parsimonious theoretical account of retrieval. Discussed in more detail in the General Discussion, this interpretation simply construes mean recall latency as a measure of the breadth of search. That is, an increase in the *search set* (which includes the internal representations of target items and possibly extralist items) results in slower retrieval, on the average, of a particular item within the search set. This assumption is consistent with the results of a recent study of proactive interference by Wixted and Rohrer (1993). In that experiment, subjects underwent three successive Brown-Peterson trials with different study lists drawn from the same category followed by a single list of items belonging to a new category. As expected, recall probability declined across the first three trials and rebounded on the fourth. More notably, with each new list of words belonging to the original category—that is, with the accumulation of previously studied words categorically related to those to be retrieved—latencies slowed. Likewise, after study of the last list of items, which were categorically unrelated to all previous items, latencies quickened significantly. Thus, with the reasonable and widely held assumption that search sets include items from the most recent list as well as categorically related items from prior lists, these effects of proactive interference on latency are to be expected (cf. Gorfein & Jacobson, 1973).

The four experiments in the present investigation were rather straightforward: a single-subject analysis, a manipulation of list length, a manipulation of study duration, and a manipulation of both list length and study duration. The analyses were threefold. First, the rela-

tionship between probability of recall and latency to recall was examined, as described above. Second, latency distributions were tabulated and compared to a theoretical distribution that provides an interesting account of retrieval. Third, the growth of IRTs as a function of output position was examined. Each of these analyses tells the same story. Before presentation of the experiments, latency distributions and IRT growth will first be discussed in more detail.

Latency Distributions

There are a few methodological differences between the present analysis of latency distributions and those carried out by previous investigators (cf. Bousfield, Sedgewick, & Cohen, 1954; Gronlund & Shiffrin, 1986; Roediger, Stellon, & Tulving, 1977; Roediger & Thorpe, 1978; Roediger & Tulving, 1979). (The results of these latency studies and previous IRT studies, for both episodic and semantic memory, are reviewed in Wixted & Rohrer, 1994.) First, these researchers, who relied on written recall, grouped latencies into relatively large bins, generally 1 or 2 min in width. In the present study, the use of a computer and a voice key magnified the temporal resolution and permitted the analysis of the critical early moments of recall. Second, whereas previous researchers have tabulated cumulative latency distributions, the total number of items recalled prior to each point in time, we present noncumulative latency distributions, as illustrated in Figure 1. (This figure includes data from Experiment 1 and is used here for illustrative purposes.) Although the cumulative curve is striking, its smoothness results from the dependence of each data point upon the previous data point. The noncumulative distribution, though noisier, is more telling, especially in its depiction of the pause that precedes free recall. However, both the cumulative and the noncumulative distributions, when fit by mathematical functions, divulge the same information about the rate of retrieval.

As illustrated in Figure 1, and reported by Roediger et al. (1977), cumulative latency distributions are described well by the cumulative exponential,

$$R(t) = N\left(1 - e^{-t/\tau}\right),$$

where $R(t)$ equals the cumulative number of items recalled by time t, N represents the number of items recalled after infinite time (i.e., asymptotic recall) and τ represents the mean latency of the recalled items. In fact, the cumulative exponential in Figure 1 accounts for 99.8% of the variance. Notice, however, that this curve is right shifted to account for the pause that precedes recall, thereby indicating that the free recall process includes a brief initiation stage followed by a single, ongoing search process that yields exponential cumulative recall. Thus, we would expect noncumulative latency distributions to be described well by the theoretical distribution that describes just such a process: the ex-Gaussian (cf. Burbeck & Luce, 1982; Dawson, 1988;

Heathcote, Popiel, & Mewhort, 1991; Hockley, 1982, 1984; Hohle, 1965; Luce, 1986; Ratcliff, 1978, 1979; Ratcliff & Murdock, 1976). Indeed, as illustrated in Figure 1 and reported by Wixted and Rohrer (1993), free recall latencies are described well by the ex-Gaussian distribution,

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

as derived in Appendix A. Although the ex-Gaussian is rather unsightly, it is conceptually quite elegant. Specifically, ex-Gaussian latencies result when a response process consists of two serial, independent stages, a normally (or Gaussian) distributed stage and an exponentially distributed stage. The duration of the normal stage has mean μ (mu) and standard deviation σ (sigma), and the duration of the exponential stage is solely described by τ (tau), which equals both the mean and the standard deviation. Most importantly, the sum of μ and τ necessarily equals mean latency. With regard to free recall, the normal stage of the ex-Gaussian presumably represents a brief initiation that precedes the retrieval of the first item and the exponential stage represents an ongoing search that yields target items at a rate that declines ex-
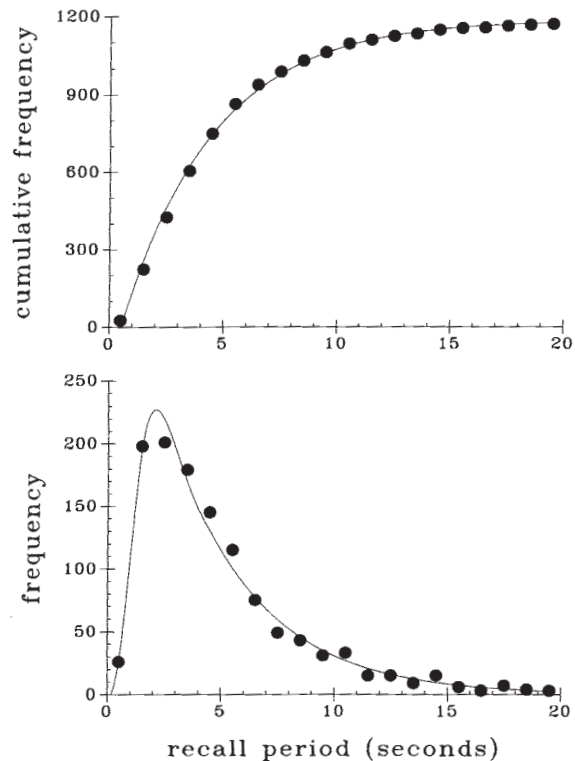

Figure 1. Recall latency cumulative distribution with best-fitting cumulative exponential distribution, and recall latency (noncumulative) distribution with best-fitting ex-Gaussian distribution (Experiment 1).

ponentially (or, as described above, accumulates exponentially, as derived in Appendix B).

The initiation that precedes free recall lasts about a second and includes the perception of the recall prompt and perhaps the delimiting of the search set. Admittedly, this initiation stage might be described equally well by other unimodal distributions such as the lognormal, gamma, generalized gamma, or Maxwell. One disadvantage of the normal distribution is its domain of both negative and positive values, thereby theoretically allowing for both negative and positive reaction times. However, this drawback is of no practical concern if the normally distributed stage of the ex-Gaussian lies almost entirely to the right of the origin—that is, when μ is greater than 2σ.

There are, of course, theoretical advantages to the normal distribution. For instance, if two or more serial stages of a response process each have durations that are normally distributed, the sum of those durations is itself normally distributed. Moreover, if a response process includes roughly 10 or more consecutive serial stages, distributed normally or otherwise, the total response time approaches normality, given that no one stage contributes disproportionately to the variance of the sum (Hohle, 1965). Thus, the normal distribution seems appropriate for the initiation stage, for this stage may itself comprise multiple substages. But the theoretical properties of the normal are not critical in this analysis. Essentially, this brief stage is needed to account for the short but distinct pause that precedes the recall of items.

Less arbitrarily, the search stage of free recall seems uniquely well described by the exponential. Evidence that this search process is exponential derives from two sources. First, the exponential growth of cumulative recall (discounting the initiation) described above necessitates that the rate of recall declines exponentially (see Appendix B). Second, the tails of recall latency distributions decline exponentially (Wixted & Rohrer, 1993), indicating that the response process includes an exponential stage whose duration constitutes a large proportion of the total response time (Ashby, 1982).

Exponentially declining rates of recall (i.e., latency distributions with exponential tails) are characteristic of the *pure-death process* (McGill, 1963). According to this simple model of retrieval, each item within a set has the same probability of retrieval (or "death"), and this probability holds constant throughout the retrieval period. Once an item has been retrieved, subsequent retrievals of that same item are ignored, and the rate of recall consequently declines. Search continues until all items have been retrieved. Note that the probability of retrieval is unaffected by either prior events or the passage of time, and, for this reason, the exponential is often characterized as "memoryless."

The most widely known instantiation of the pure-death process is that of random sampling (with replacement) (see Albert, 1968; Herrmann & Pearle, 1981;

Indow & Togano, 1970; McGill, 1963; Murdock & Okada, 1970; Schulz & Albert, 1976; Shiffrin, 1970; Vorberg & Ulrich, 1987). According to this serial interpretation, which is derived in Appendix B, items are randomly sampled one at a time, at a constant rate, from a search set. This search set includes target items and, possibly, extralist items. Since it is generally believed that the representation of a list item (i.e., a memory trace) can vary in strength to the extent that retrieval may or may not be possible, we will assume further that the search set includes both recoverable and nonrecoverable target items. After sampling, each item is immediately recognized as either a yet-to-be-retrieved recoverable target (and then recalled) or a previously retrieved target, nonrecoverable target, or extralist item (and then ignored). As the number of yet-to-be-retrieved recoverable targets decreases, the rate of recall correspondingly declines. As discussed in more detail in the General Discussion, both this account and a corresponding parallel interpretation dictate that mean latency is directly proportional to the number of items within the search set. Thus, the use of latency as a measure of search set breadth is suggested both by intuition and by the exponential nature of search.

Independent of its theoretical implications, an ex-Gaussian latency distribution results from a response process that includes two distinct stages. This observation is important, because experimental manipulations may differentially affect the duration of these two stages. Such dissociations of the ex-Gaussian parameters μ and τ have been found in simple reaction tasks (Hohle, 1965), recognition memory (Hockley, 1982; Ratcliff & Murdock, 1976), visual search and working memory (Hockley, 1984), and the Stroop effect (Heathcote et al., 1991). Dissociations of μ and τ are reported in the present study as well.

IRT Growth

IRTs may provide the most accurate portrayal of the time course of free recall once the search process is underway. In the present study, we examine the growth of mean IRTs as a function of output position, as was done in each of the previous episodic IRT studies (Murdock & Okada, 1970; Patterson, Meltzer, & Mandler, 1971; Pollio, Kasschau, & DeNise, 1968; Pollio, Richards, & Lucas, 1969). Each of these studies employed an immediate recall procedure, thereby ensuring the contribution of short-term storage as well as long-term storage (Raaijmakers & Shiffrin, 1980). It therefore remains unclear whether the results of these IRT studies would differ with the use of a filled retention interval. More definitively, subjects in the Murdock and Okada study recalled 7 or fewer items from a 20-item study list in three fourths of the trials, suggesting that the long-term store contributed only a small proportion of the responses. Metcalfe and Murdock (1981) reach the same conclusion after analyzing these data further. It is of note

that the study by Murdock and Okada is the only previous one in which the mathematical form of IRT growth has been examined.

Intriguingly, though the pure-death process account of retrieval yields both exponentially increasing cumulative recall and an exponentially decreasing rate of recall, it does *not* yield exponential IRT growth. In fact, the pure-death process actually dictates hyperbolic IRT growth. According to this equation, hereafter referred to as the *pure-death hyperbola*, the

$$\text{Mean } i\text{th IRT } = \frac{\tau}{N-i} \text{ for } i = 1, 2, \ldots, N-1$$

where τ, the only parameter, again represents mean latency, N equals the recall total for that trial, and i equals the ordinal position of the IRT so that the 1st IRT occurs between the recall of the first and second items. (If there were no initiation stage preceding the search stage, the substitution of 0 for i in this equation would give the duration of the so-called null IRT, the latency of the first item.) This equation is derived in Appendix C.

This equation reflects an important characteristic of IRT growth in a pure-death process. Namely, for a given search-set size (i.e., a given value of τ), an IRT is uniquely determined by the number of items not yet recalled $(N-i)$. Thus, the last IRT of a four-item recall should equal the last IRT of a nine-item recall from the same sized search set. In particular, the last IRT equals $\tau/1$, the second to last IRT equals $\tau/2$, etc., regardless of recall total (N). Therefore, the growth of mean IRTs given by a pure-death process can be *collapsed* across recall total by "reverse output position" (last IRT, second-to-last IRT, etc.). In this way, a single estimate of τ can be obtained for each experimental condition.

Finally, we should point out that the analyses of IRT growth and latency distributions are complementary rather than redundant. For example, the pause preceding the recall of the first item is reflected in the latency distribution but excluded by an analysis of IRT growth. By the same token, but more importantly, estimates of τ via IRTs serve to confirm estimates of τ via latencies, since the parameter estimates given by the ex-Gaussian can be misleading. In particular, it is possible that the ex-Gaussian nature of a particular latency distribution is only an artifact of summing latencies across trials in which recall totals vary. In fact, even a constant rate of recall can theoretically yield ex-Gaussian distributions after summing across trials. Fortunately, the estimates of τ via IRTs can rule out such possibilities.

GENERAL STATISTICAL TECHNIQUE

Before all data analyses, response latencies resulting from voice key false alarms (i.e., activation of voice key by extraneous noises) were removed. For each voice key miss, the entire trial including the miss was excluded from the data analysis in order to avoid significant distortion of mean IRTs. Fortunately, voice key misses were rare (less than one per hundred hits), since voice key sensitivity was set relatively high. That is, the number of voice key misses was minimized at the expense of more voice key false alarms, because the latter can be removed without distorting the data.

False alarms by subjects were also rare, occurring at a rate of less than 2% in each experiment. Since the removal of false alarms distorts the adjacent IRTs, we decided a priori to include any false alarms in our analyses, because temporal measures were our primary concern. However, the removal of so few would not affect our basic findings.

The mean recall latencies given in Experiments 2, 3, and 4 are means of subjects' means. (Experiment 1 is a single-subject analysis.) In contrast, the latency distributions and the mean IRTs reported in this paper were derived by summing across all *responses*. Thus, the mean recall latencies differ slightly from the sum of μ and τ, because the number of items recalled by each subject varied.

Observed latency distributions were fit by the ex-Gaussian distribution according to a maximum likelihood estimation procedure (cf. Maindonald, 1984; Ratcliff & Murdock, 1976). A chi-square goodness-of-fit statistic was calculated for each best-fitting ex-Gaussian. As is done typically, bins with few responses were combined (and degrees of freedom reduced) to yield expected values of at least five for each bin (cf. Hoel, 1971; Ratcliff & Murdock, 1976). Further, in order to ensure that the normal stage of the ex-Gaussian lay primarily to the right of the origin (see the introduction), it was confirmed that μ was greater than 2σ in each fit. Finally, t tests of significant difference were performed on pairs of parameter estimates by obtaining asymptotic standard errors (ASEs) from the Hessian matrix of second partial derivatives (cf. Maindonald, 1984; Ratcliff & Murdock, 1976).

The IRTs reported for the following experiments measured the duration between the voice onsets of consecutively recalled items rather than the duration of the "downtime" between an item's voice offset and the next item's voice onset. Though both methods have been used before, we prefer, as do Patterson et al. (1971), not to ignore the vocalization time. Not only might the vocalization of an item overlap with the internal search for the next item, the exclusion of vocalization time would theoretically disallow the comparison of results given by IRTs and latencies, since the latter necessarily include the vocal durations of items recalled earlier in the trial. In any case, our exclusive use of four-letter monosyllabic study words minimized any possible effects of including vocalization time.

The best-fitting pure-death hyperbolas for each plot of IRT growth were determined by least squares. The growth of mean IRTs as a function of serial position was plotted in two ways: either by partitioning the data by the total number of items recalled on each trial (Figures 3 and 6)

or by collapsing across reverse output position (Figures 4 and 7), as described in the introduction. The plots of IRT growth partitioned by recall total exclude recall totals that occurred less than five times as well as recall totals of two, because a single IRT cannot convey IRT growth. However, the plots of IRT growth collapsed across reverse output position include all IRTs. Likewise, all IRTs were included when fitting the pure-death hyperbola.

EXPERIMENT 1

To ensure that the ex-Gaussian nature of latency distributions is not an artifact of pooling data across subjects, we performed a preliminary experiment that involved eight single-subject analyses.

Method

Subjects. Eight undergraduates of the University of California, San Diego participated for course credit.

Materials. Study-list items were randomly chosen from a pool of four-letter monosyllabic nouns, and their selection was randomized uniquely for each subject.

Design. Each study list contained five words, which were presented once every 2 sec. After three practice trials, each subject completed two 25-trial sessions.

Equipment. A Lafayette Instruments Voice Activated Relay (Model 18010), in conjunction with an IBM-compatible personal computer, recorded the onset of each spoken word to the nearest millisecond.

Procedure. Subjects were tested in the presence of an experimenter, and the trials were computer driven. In the study period, beginning after a 3-sec prompt, a new word appeared every 2 sec (and remained throughout the study period), and subjects read the words aloud as they appeared. In the distractor task, 10 three-digit numbers appeared for 2 sec each, and the subjects read the three digits aloud in ascending order of value. In the 20-sec recall period, the subjects recalled the study words aloud and an experimenter monitored the recall latencies via computer in order to subject false alarms as well as voice key false alarms (e.g., the recording of a cough) and voice key misses (e.g., the failure to record a soft-spoken response). After a 10-sec rest period, the next trial began.

Results and Discussion

Table 1 includes mean recall probability, mean recall latency, and the best-fitting ex-Gaussian latency distribution for each of the 8 subjects. For all but Subject 3, the best-fitting ex-Gaussian produced nonsignificant chi-square goodness-of-fit statistics, indicating an ac-

ceptable fit. The anomalous data of Subject 3 were most probably due to a ceiling effect; she recalled all five items in 82% of the trials whereas the other 7 subjects, on the average, recalled all five items in only 20% of the trials. The correlation between recall probability and recall latency was $-.39$ with Subject 3 included and $+.09$ with Subject 3 excluded, both of which were nonsignificant. The recall latencies, collapsed across all subjects except Subject 3, are those used in Figure 1 to illustrate cumulative and noncumulative latency distributions.

EXPERIMENT 2

In this experiment, study-list length was manipulated. Although a longer study list does lead to longer recognition latencies (see, e.g., Atkinson & Juola, 1974), its effect on recall latency is unknown. However, an examination of list-length effects on cumulative recall by Roediger and Tulving (1979) suggests that free recall mean latency is affected in the same manner as recognition latency is. Although Roediger and Tulving did not focus on either latency or the mathematical form of cumulative recall, they noted that subjects recalling short lists "reached asymptotic level of recall at a greater rate" (p. 611). Because faster approaches to the asymptote generally translate into faster mean latencies, it is probable that the longer list lengths resulted in reduced mean latencies in that experiment.

Such a finding can easily be explained by a search-set analysis. Given that longer study lists result in search sets that contain more targets (and perhaps more extra-list items), mean recall latency should increase. More precisely, estimates of the initiation time (μ) should remain constant, while estimates of the search time (τ), given by either latency distributions or IRT growth, should increase.

Incidentally, although shorter study lists yield the recall of fewer list items (but a greater proportion of list items) than longer lists do, it is not the case that any manipulation that reduces recall total will necessarily *shorten* mean latencies. For instance, the buildup of proactive interference, as described in the introduction, results in fewer recalled items (and reduced recall probability) but *longer* mean latencies. Moreover, if the recall trials within any particular condition are grouped by the number of items recalled (N), mean latency is independent of N.

Method

Subjects. Twelve undergraduates of the University of California, San Diego participated for course credit.

Materials. Study-list items were randomly chosen from a pool of four-letter monosyllabic nouns, and their selection was randomized uniquely for each subject.

Design. Each study list contained three, six, or nine words, which were presented once every 2 sec. After three practice trials, 24 scored trials (8 of each length) appeared in an order randomized uniquely for each subject.

Equipment. A Lafayette Instruments Voice Activated Relay (Model 18010), in conjunction with an IBM-compatible personal

Table 1
Mean Recall Probability, Mean Recall Latency (in Seconds), and Latency Distributions in Experiment 1

Subject	Probability	Latency	Ex-Gaussian Fits of Latency Distributions		
			χ^2	(df)	p
1	.67	4.07	3.98	(6)	.68
2	.62	5.05	11.42	(7)	.12
3	.94	3.45	27.21	(6)	<.01
4	.60	4.90	3.82	(6)	.70
5	.62	6.09	6.13	(9)	.73
6	.62	3.94	4.98	(6)	.55
7	.88	5.19	5.02	(8)	.76
8	.68	4.54	3.66	(6)	.72

Table 2
Mean Recall Probability (With Standard Errors) and Mean Recall Latency (in Seconds, With Standard Errors) in Experiment 2

List Length	Probability		Latency	
	M	SE	M	SE
3 words	.85	.02	3.22	.23
6 words	.66	.03	5.77	.28
9 words	.52	.04	6.90	.40

Table 3
Fits of Latency Distributions (Figure 2) and Interresponse Time Growth (Figure 4) in Experiment 2

List Length	Ex-Gaussian Fits of Latency Distributions					Pure-Death Fits of IRT Growth		
	χ^2	(df)	μ	SE	τ	SE	VAF	τ	SE
3 words	9.16	(5)	0.90	.12	2.25	.18	.77	2.53	.19
6 words	16.74	(12)	1.16	.11	4.66	.28	.90	4.56	.22
9 words	24.89	(13)	1.07	.13	6.19	.36	.91	5.97	.26

Note—Parameter estimates are given in seconds and are accompanied by their asymptotic standard errors. Each χ^2 statistic was nonsignificant ($p > .1$).

computer, recorded the onset of each spoken word to the nearest millisecond.

Procedure. Subjects were tested in the presence of an experimenter, and the trials were computer driven. In the study period, beginning after a 3-sec prompt, a new word appeared every 2 sec (and remained throughout the study period), and subjects read the words aloud as they appeared. In the distractor task, 10 three-digit numbers appeared for 2 sec each, and the subjects read the three digits aloud in ascending order of value. In the 30-sec recall period, the subjects recalled the study words aloud and an experimenter monitored the recall latencies via computer in order to record subject false alarms as well as voice key false alarms (e.g., the recording of a cough) and voice key misses (e.g., the failure to record a soft-spoken response). After a 10-sec rest period, the next trial began.

Results and Discussion

As can be seen in Table 2, longer study lists resulted in decreased recall probability and increased recall latency, as expected. Analyses of variance indicated a significant decrease in recall probability [$F(2,22) = 46.17$, $MS_e = .01$] and a significant increase in recall latency [$F(2,22) = 67.47$, $MS_e = .64$], as list length increased. (Throughout this paper, any test described as significant produced p values less than .01, and any test described as insignificant produced $p > .1$.)

As can be seen in Figure 2 and Table 3, the latency distributions were well described by the ex-Gaussian. Estimates of μ, each about 1 sec, varied nonmonotonically across increasing list lengths, as is seen in the similarity in the ascending arms of each distributions. Estimates of τ, on the other hand, grew from about 3 sec to 7 sec, as is illustrated by the differential declines in the tails of each distribution. Each of the three pairwise comparisons of μ were insignificant ($p > .1$), whereas the three pairwise comparisons of τ were significant.

As can be seen in Figure 3, the growth of mean IRTs, partitioned by recall totals, is described well by the one-

parameter pure-death hyperbola. Likewise, Figure 4 nicely depicts the pure-death hyperbolic growth of mean IRTs that are collapsed across recall total and plotted as a function of reverse output position (e.g., $N-1$ denotes the last IRT). These fits did, however, underestimate the "least recent" IRTs in each condition, due, in part, to the rarity of the longer recall totals that provided these "least recent IRTs." For example, in the nine-study-words condition, eight or more words were recalled only five times, with four of those occurrences given by a single subject. Thus, the second data point in the nine-study-words IRT growth plot (i.e., the $N-7$ interval) is the average of only five IRTs. Nevertheless, the fits in Figure 4 accounted for a large portion of the variance, as reported in Table 3. Further, these fits yielded estimates of τ that increased with study list length, yielding three significant pairwise comparisons. Finally, these three estimates of τ via IRTs were not significantly different from the corresponding estimates of τ via latencies in each condition.

The random sampling interpretation of the pure-death process offers one explanation of the effects of study-list length on mean latency. First, longer lists result in a greater number of target items (and possibly extralist items) within the search set, which in turn increases mean latency. Second, longer study lists result in an in-
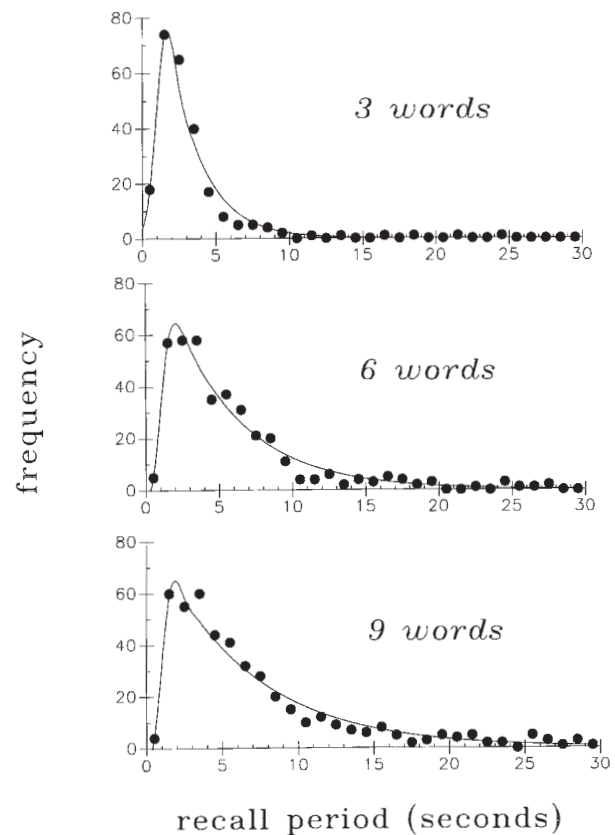
Figure 2. Recall latency distributions, grouped into 1-sec bins, with best-fitting ex-Gaussians for varying study list lengths (Experiment 2).
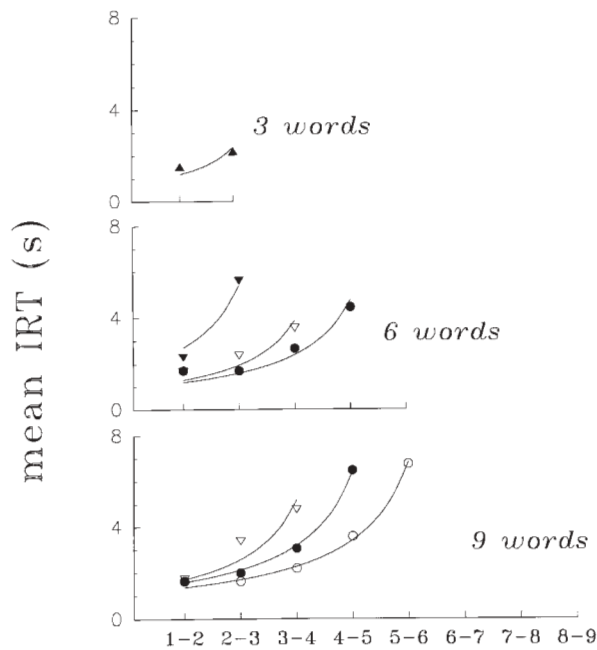
Figure 3. The growth of mean interresponse times (IRTs), partitioned by recall total, with best-fitting pure-death hyperbola for varying list lengths (Experiment 2).
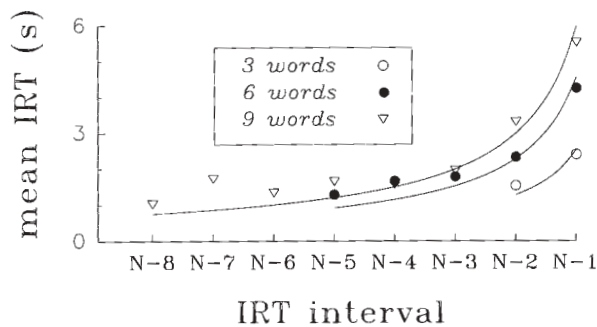

Figure 4. The growth of mean interresponse times (IRTs), collapsed across recall total, with best-fitting pure-death hyperbola for varying list lengths (Experiment 2). The mean IRTs are plotted as a function of reverse output position, where $N-1$ denotes the last IRT.

crease in the number of recoverable targets within the search set, as reflected in the greater *absolute* number of items recalled. However, the increase in the number of recoverable target items is less than the increase in the number of study-list items. Thus, the *proportion* of list items that are recoverable (i.e., recall probability) declines with longer list lengths.

EXPERIMENT 3

In this experiment, item study duration was manipulated. While longer exposure to each list item would certainly increase recall probability, its effect on recall latencies was again unknown. Once again, though, a pre-

vious finding suggested a possible outcome. Bousfield et al. (1954) presented study lists to subjects one, two, three, four, or five times and then plotted cumulative recall across five 2-min intervals. We visually estimated these data points, and, by incorporating the curve fits presented by Bousfield et al., we were able to derive estimates of mean latency for each condition. The increase in the number of list presentations resulted in a slight increase in mean latency. This result seems to us rather striking, because intuition might suggest that items that are learned better are recalled more quickly. Given that a greater number of list presentations results in greater study time for each item, we might expect (on the basis of the data from Bousfield et al.) that the manipulation of item study duration would not decrease mean latency.

Method

In Experiment 3, six-word study lists were presented to subjects at rates of 1, 2, or 4 sec per word. Each subject received seven trials per condition. Otherwise, Experiment 3 was identical to Experiment 2.

Results and Discussion

As can be seen in Table 4, the increase in study item exposure resulted in higher recall probability, as expected, but had no effect on mean recall latency. Likewise, analyses of variance indicated a significant increase in recall probability [$F(2,22) = 42.73$, $MS_e = .01$] and insignificant differences in recall latency [$F(2,22) < 1$].

As can be seen in Figure 5 and Table 5, the ex-Gaussian described the latency distributions well. The differences between each of the three pairs of τ values were not significant, as is illustrated by the similarity among the three distributions. Unexpectedly, estimates of μ increased from .86 to 1.36 sec with greater exposure, though the difference between these two extreme values was not significant. Notably, this increase in μ was seen in another experiment in our laboratory in which study lists were presented one, two, or four times at a constant rate. Thus, it seems that longer exposures by way of either slower presentation rates or multiple list presentations produce a longer initiation stage. However, the effects are small, and the result should be considered tentative.

Figure 6 shows the pure-death hyperbolic growth in the mean IRTs and clearly illustrates the dependence of IRTs on the number of yet-to-be-recalled items. The last IRTs of each recall total are between 4 and 5 sec, the second to last IRTs of each recall total are between 2 and

Table 4
Mean Recall Probability (With Standard Errors) and
Mean Recall Latency (in Seconds, With Standard Errors)
in Experiment 3

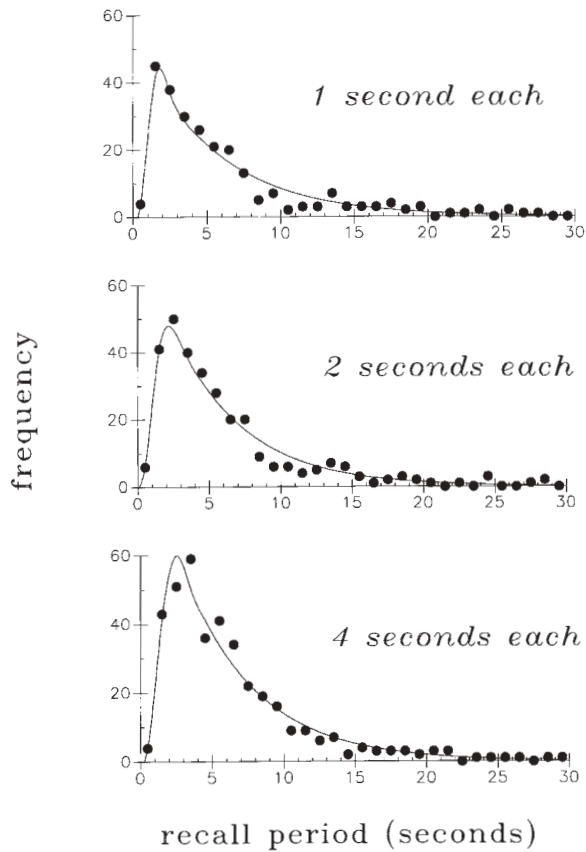| Study	Probability		Latency	
Exposure	M	SE	M	SE
1 sec each	.50	.03	6.07	.41
2 sec each	.60	.03	5.96	.40
4 sec each	.76	.03	6.30	.43

Figure 5. Recall latency distributions, grouped into 1-sec bins, with best-fitting ex-Gaussians for varying study item exposure (Experiment 3).

Table 5
Fits of Latency Distributions (Figure 5) and Interresponse Time Growth (Figure 7) in Experiment 3

Study Exposure	Ex-Gaussian Fits of Latency Distributions						Pure-Death Fits of IRT Growth		
	χ^2	(df)	μ	SE	τ	SE	VAF	τ	SE
1 sec each	17.10	(12)	0.86	.20	5.30	.39	.96	5.39	.23
2 sec each	12.33	(13)	1.15	.13	4.89	.33	.98	5.21	.13
4 sec each	12.65	(14)	1.36	.14	5.03	.31	.97	4.96	.15

Note—Parameter estimates are given in seconds and are accompanied by their asymptotic standard errors. Each χ^2 statistic was nonsignificant ($p > .1$).

3 sec, and the third to last IRTs are less than 2 sec, regardless of recall total. Such parity is expected in a pure-death process.

Consequently, the plots of mean IRTs in Figure 7 that are collapsed across recall total by reverse output position are fit well by the pure-death hyperbola, accounting for more than 95% of the variance, as reported in Table 5. The similarity between these three plots provides further evidence that mean latency is unaffected by the duration of study item exposure. As expected, the estimates of τ given by these three fits, as reported in Table 5, yielded three nonsignificant pairwise comparisons. Finally, as in the last experiment, these estimates

of τ were not significantly different from the corresponding estimates of τ via latencies in each condition (also reported in Table 5).

The random sampling interpretation of the pure-death process again provides a mechanism by which to understand the effect of varying study durations upon recall probability and latency. The increase in study duration resulted in a greater number of recoverable target items within the search set, thereby increasing recall probability. However, the total number of recoverable and non-recoverable targets (and extralist items, if any) remains constant with longer study duration, thereby not affecting mean latency. Thus, longer study times do not result in a greater number of targets within the search set; it instead increases the likelihood that a target item within the search set will be recoverable. This is not the only possible interpretation of this result, but it is the one most easily reconciled with the results of Experiments 1 and 2 and the recall latency analyses of proactive interference described earlier.

EXPERIMENT 4

Thus far, we have obtained both a positive correlation and a noncorrelation between recall probability and recall latency. Taken together, these two findings represent
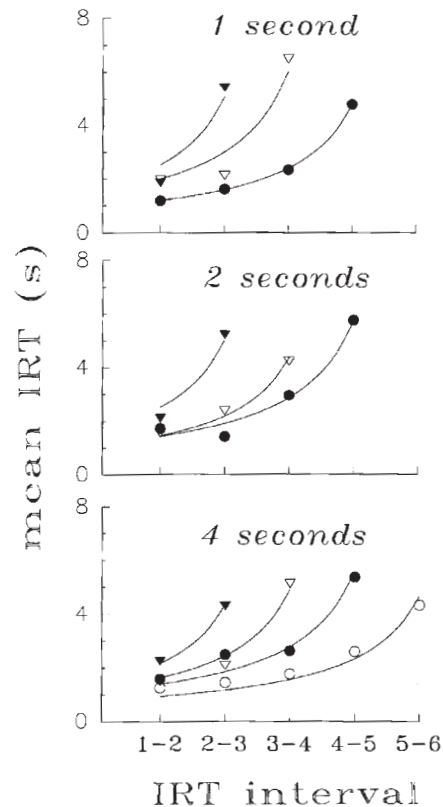

Figure 6. The growth of mean interresponse times (IRTs), partitioned by recall total, with best-fitting pure-death hyperbola for varying study item exposure (Experiment 3).
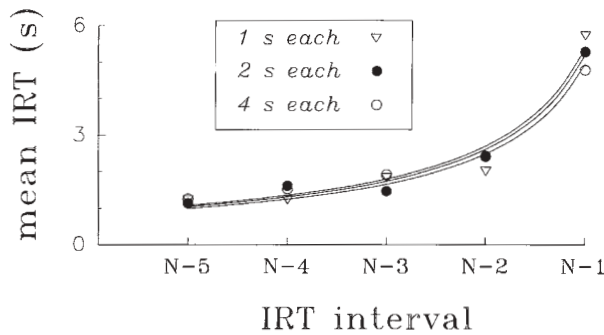
Figure 7. The growth of mean interresponse times (IRTs), collapsed across recall total, with best-fitting pure-death hyperbola for varying study item exposure (Experiment 3). The mean IRTs are plotted as a function of reverse output position, where $N-1$ denotes the last IRT.

a dissociation of the two dependent measures; however, an inherent relationship can be completely ruled out. As MacLeod and Nelson (1984) argue in their dissociation of probability and latency in cued recall, two measures that are a priori associated may sometimes appear non-correlated if a manipulation does not vary across a wide enough range of values. If, for example, presentation rate in the last experiment varied by a factor of 10 rather than 4, latency might then have been affected.

The results of the first two experiments suggest an experimental manipulation that addresses the issue in a more definitive way. Namely, if study lists were both lengthened *and* presented more slowly, then mean latency should increase (due to the longer lists) and recall probability should also increase (due to the greater study time). In fact, since the results of the first two experiments dictate this finding, it might have been presented as just a thought experiment. But in the interest of empiricism, we present this extension (as well as replication) of Experiments 2 and 3 and a brief analysis.

Method

In Experiment 4, subjects were presented with either six-word lists at 1 sec per word or nine-word lists at 4 sec per word. Because Experiment 3 included just two conditions, one third fewer subjects were tested than in Experiments 2 and 3. Otherwise, the method was identical to that described for Experiment 2.

Results and Discussion

The shift from the six quickly presented items to the nine slowly presented items produced a highly significant increase in both recall probability [$F(1,7) = 12.50$, $MS_e = .01$] and mean recall latency [$F(1,7) = 30.86$, $MS_e = .82$], as is shown in Table 6.

The ex-Gaussian fits to both latency distributions yielded nonsignificant chi-square statistics, as is reported in Table 7. The two estimates of μ were not significantly different (though there was once again a perceptible increase with longer study exposure) and the two estimates of τ were significantly different.

Pure-death hyperbolic fits of mean IRT collapsed across recall total and plotted with respect to reverse out-

put position were noisier than those in Experiments 2 and 3, yet nevertheless accounted for most of the variance, as reported in Table 7. These fits yielded two estimates of τ that were not quite significantly different, though both estimates of τ via IRTs were not significantly different from the corresponding estimates of τ via latencies.

In sum, the manipulation of both list length and presentation rate resulted in a positive correlation between recall probability and latency, and the effect on mean latency stems primarily from the variability in the duration of retrieval (τ), not initiation (μ). As interpreted by a search-set analysis, a longer, slowly presented list results in a search set that includes both more items (and therefore longer mean latencies) and a greater number of recoverable targets (and therefore better recall probability).

GENERAL DISCUSSION

We have reported a single-subject analysis of recall latency and three experiments that demonstrate a negative, nonexistent, and positive correlation between mean recall probability and mean recall latency. In Experiment 2, an increase in list length decreased recall probability and increased recall latency. In Experiment 3, an increase in study item exposure increased recall probability but did not affect recall latency. In Experiment 4, an increase in both list length and study item exposure increased both recall probability and recall latency.

In addition, the ex-Gaussian nature of latency distributions, whether given by single subjects (Experiment 1) or summed across subjects (Experiments 2, 3, and 4), and the independent manipulation of its parameters μ and τ suggest that the retrieval process consists of a brief, normally distributed initiation stage followed

Table 6
Mean Recall Probability (With Standard Errors) and
Mean Recall Latency (in Seconds, With Standard Errors)
in Experiment 4

Study Condition	Probability		Latency	
	M	SE	M	SE
6 words at 1 sec each	.40	.04	4.87	.74
9 words at 4 sec each	.56	.04	7.38	.85

Table 7
Fits of Latency Distributions and Interresponse Time
Growth in Experiment 4

Study Condition	Ex-Gaussian Fits of Latency Distributions						Pure-Death Fits of IRT Growth		
	χ^2	(df)	μ	SE	τ	SE	VAF	τ	SE
6 words at 1 sec each	16.17	(10)	0.72	.14	4.33	.30	.55	4.62	.39
9 words at 4 sec each	22.81	(19)	0.95	.16	6.74	.38	.93	5.93	.23

Note—Parameter estimates are given in seconds and are accompanied by their asymptotic standard errors. Each χ^2 statistic was nonsignificant ($p > .1$).

by an exponentially distributed search stage. Further, pure-death hyperbolic fits of the IRT growth provided estimates of τ that were comparable to those given by the ex-Gaussian fits of latency distributions. Finally, both ex-Gaussian latency distributions and pure-death hyperbolic IRT growth are consistent with a pure-death process model of retrieval, which is now discussed in more detail.

Theoretical Interpretations

The interpretation of latency as a measure of the breadth of search is not only intuitive, it is necessitated by the simplest theoretical account of exponentially declining retrieval, the random-sampling-with-replacement scheme. The serial interpretation of this account is well known (see Albert, 1968; Indow & Togano, 1970; McGill, 1963; Murdock & Okada, 1970; Schulz & Albert, 1976; Shiffrin, 1970; Vorberg & Ulrich, 1987), but a closely related parallel version exists as well.

The serial interpretation was presented earlier. Briefly, items within the search set are randomly selected, one at a time, at a constant rate. Selected items are immediately recognized as either a not-yet-recalled target item (and then recalled) or a previously recalled target, nonrecoverable target, or extralist item (and then ignored). As the number of not-yet-recalled target items decreases, the rate of retrieval consequently declines.

Interpreted in parallel, all items within the search set have an equal momentary probability of retrieval that remains constant until an item is retrieved, after which the momentary probability of retrieval equals zero. In addition, a finite amount of resources is distributed equally among each item within the search set, and the retrieval probability for each item is a direct function of the amount of resources allocated for that item. Therefore, if the number of items within the search set is, say, doubled, the allocation of resources for each item, which directly affects retrieval probability, is consequently cut in half. Unfortunately, the serial and parallel interpretations of the pure-death process, as described above, cannot be teased apart experimentally (cf. Townsend, 1971; Vorberg & Ulrich, 1987).

Most importantly, the random sampling model offers insight into the size of the search set. Very simply, as derived in Appendix B, the parameter estimate $\tau = St^*$, where S equals the number of items within the search set and t^* represents the shortest possible retrieval time—that is, the time to retrieve an item when the search set contains only that one item. Thus, in the serial version, t^* equals the duration of a single sampling, and in the parallel version, t^* equals the expected time to retrieve an item when all resources are focused on the retrieval of that single item.

Of course it is not possible to accurately measure sampling time (t^*), but given that it remains roughly constant, one can make relative comparisons of search-set size (S) between conditions based on values of τ. For instance, if an experimental manipulation results in the doubling of τ, the model would suggest that the number of items within the mental search set doubled as well. Thus, this search-set sampling account implies what intuition suggests: mean latency reflects the size of the search set. Although the exponential form of retrieval was observed long ago, the fact that τ provides a theoretical estimate of search-set size has never been exploited.

The pure-death process, as described above, is the most straightforward version; further assumptions can be added. For example, Vorberg and Ulrich (1987) have presented a generalization of the pure-death process in which the time to sample and identify different items from the search set is not constant, but instead varies exponentially. The allowance for varied sampling rates is, of course, but one of many such elaborations that are possible. One particularly detailed extension of the pure-death process is the well-known search of associative memory (SAM).

SAM

The random sampling account of a pure-death process was used by Shiffrin (1970), and it provided the framework for the SAM model described by Raaijmakers and Shiffrin (1980). As these authors note, SAM depicts retrieval as a repeated sampling of a "search set" that contains images with differential "associative strengths" such that the retrieval probability for each item is a direct function of its associate strength divided by the sum of the associative strengths of all items. Thus, if the associate strengths of individual items does not vary considerably, the momentary retrieval probability for each item is inversely proportional to the number of items within the search set. In other words, latency reflects the size of the search set. Of course, SAM requires further assumptions in order to account for memory phenomena other than recall latency.

Raaijmakers and Shiffrin (1980) presented SAM simulations that, according to our analyses, yield the same results as those given by the experimental data of the present paper, even though no such data were available when Raaijmakers and Shiffrin performed the simulations. Specifically, these authors presented four cumulative recall curves given by a simulation that included the manipulations of list length (10 or 40 items) and study exposure (1 or 8 sec per word). We estimated the data from their Figure 11, computed recall probability and mean latency, transformed the cumulative data into latency distributions, and then fit the exponential distribution (the second stage of the ex-Gaussian). As is shown in Table 8, the increase in list length, for either study exposure, markedly reduced recall probability while more than doubling mean latency, as in our Experiment 2. The increase in study exposure, for either list length, more than doubled recall probability, and had little effect on mean latency, as in our Experiment 3. The shift from the 10 rapidly exposed items to the 40 slowly exposed items increased both probability and mean la-

Table 8
Recall Probability, Recall Latency (in Seconds),
and Latency Distributions for SAM

Study Condition	Mean Recall Probability	Mean Recall Latency	Exponential Fits of Latency Distributions	
			χ^2 (df)	τ
10 words				
at 1 sec each	.39	15.58	.09 (0)	14.35
at 8 sec each	.82	16.10	.30 (1)	14.92
40 words				
at 1 sec each	.24	26.45	2.31 (3)	26.23
at 8 sec each	.54	34.92	6.24 (6)	37.01

Note—Each χ^2 statistic, excluding the one with 0 df, was nonsignificant ($p > .1$).

tency, as in our Experiment 4. Finally, the exponential provided decent fits of the latency distributions, as is shown in Figure 8. (Estimates of τ differed from mean latency because of the large bins, and statistical tests of their chi-square values were not possible without knowledge of the number of simulated trials.)

In the same paper, Raaijmakers and Shiffrin present simulations of IRT growth partitioned by recall total for a study list of 15 words presented for 2 sec each. These plots of IRT growth exhibit the same characteristics as those reported in this paper. In particular, for each recall total, the last IRTs, the second to last IRTs, etc., are strikingly similar in magnitude. Indeed, Raaijmakers and Shiffrin note that "at any given output position the interresponse time is a good predictor of the number of words yet to recall" (p. 235). However, after visual estimation of data points, and plotting of the growth of mean IRTs collapsed across recall totals with respect to "reverse output position," the pure-death hyperbola provided a systematically deviant fit of the simulated data. In fairness, though, a better fit might have been achieved if the mean IRTs for each recall total were first weighted by its number of occurrences before averaging the mean IRTs across recall totals (as done in Experiments 2, 3, and 4), but the necessary information was unavailable. Moreover, the analysis of IRT growth as a function of reverse output position assumes that subjects continue to
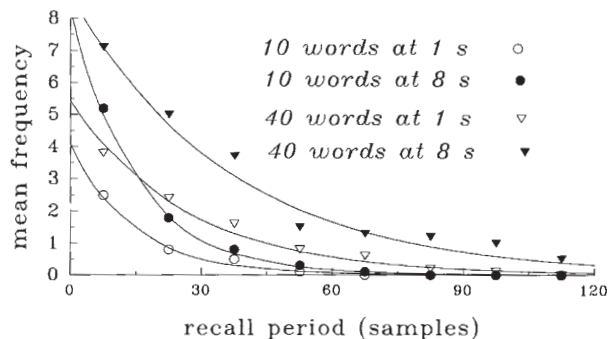
Figure 8. Recall latency distributions given by simulations of SAM and the best-fitting exponentials.

search for the "last" item within the search set, whereas SAM incorporates a stopping rule that can terminate the search process prematurely.

Though SAM includes 10 parameters, Raaijmakers and Shiffrin note that "most of the … parameters … are *not* essential for the fit of the model to most of the data" (p. 221) and that "Shiffrin (1970) fit a great deal of free recall data with just three parameters" (p. 221). Likewise, the three-parameter ex-Gaussian, a simple extension of the pure-death process, quite adequately describes the nature of retrieval.

Phenomena Not Considered

Because the pure-death process presumes that yet-to-be-recalled items have equal retrieval probabilities that are independent of each other and remain constant over time, it does not account for, *or conflict with*, several well-known experimental phenomena. First, a cluster, two or three related items recalled in quick succession, suggests that item retrievals are not always independent of each other, whereas retrievals given by a pure-death process are independent. Second, output interference, in which an item retrieval interferes with subsequent item retrievals, suggests that the item retrieval probabilities are neither independent of each other nor constant over time. Third, the pure-death process also does not explicitly address why an item's serial position in the study list affects its serial position in the output order.

Clustering. The extent of clustering depends greatly on the experimental paradigm, as revealed by our analyses. Though clustering may go undetected when latencies are summed across trials, it cannot elude an IRT analysis. Inspection of our own data revealed an occasional brief IRT (i.e., the recall of two items in quick succession) occurring relatively late in the recall chain, suggesting that the retrieval of the two items may not have been independent; that is, the two items may have formed a clustered pair. Such pairs occurred infrequently, though, for clustering was probably minimized by the use of short, categorically unrelated, once-presented study lists.

Moreover, the simulation of clustering reveals that even a moderate number of interitem associations has very little effect on the ex-Gaussian nature of latency distributions. In particular, we performed two simulations of 500 recall trials that included a brief, normally distributed initiation stage and the subsequent random sampling of an item from a search set once every second. In the first simulation, targets were not associated. In the second simulation, all targets were paired so that the retrieval of one item resulted in the retrieval of its associate in the subsequent sampling. The nonclustered simulation generated a latency distribution that was, of course, fit exceptionally well by the ex-Gaussian. Somewhat unexpectedly, the clustered simulation generated a latency distribution that was also fit well by the ex-Gaussian, as is shown in Figure 9. However, as can be seen in these simulated data, the ex-Gaussian underesti-
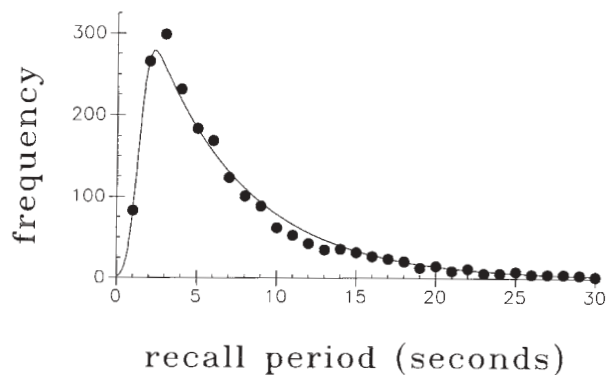
Figure 9. Recall latency distribution given by a simulation of clustering and its best-fitting ex-Gaussian distribution.

mates the upper portion of the distribution tail and overestimates the middle portion of the tail. Because items in clustered pairs have "twice the chance" of retrieval, there is an initial increase in retrieval which necessarily results in a subsequent decrease in recall. As clustering increases, this kind of deviation becomes more extreme.

The extent of clustering is most readily observed through the analysis of IRTs; in fact, of the four previous episodic IRT studies, three have focused on clustering (Patterson et al., 1971; Pollio et al., 1968; Pollio et al., 1969). In these studies, the growth of IRTs between clusters was compared with the growth of IRTs within clusters, and the results of each study were consistent with a two-stage sampling model. According to this model, a cluster is first randomly sampled from a collection of clusters, and then the items within that cluster are randomly sampled until each has been retrieved. The random selection of clusters continues until each cluster has been sampled. Thus, the growth of IRTs between clusters depends on the number of clusters, while the growth of IRTs within each cluster depends on the number of words within that cluster.

Output interference. Although it is not disputed that output interference can affect free recall, certain experimental findings that are often viewed as evidence for output interference do not conflict with the assumptions of the pure-death process. For instance, although the declining rate of recall seems to suggest that earlier recall is impeding later recall, this decline is actually predicted by the pure-death process, as can be seen in the hyperbolic growth of IRTs. Likewise, the oft-reported inhibitory effects of part-list cues upon the subsequent recall of other study items do indeed suggest output interference, but not by previously recalled items. Instead, part-list cues, which are in fact retrieval cues, may simply influence the delimiting of the search set and, consequently, the probability and latency of recall. For instance, Sloman, Bower, and Rohrer (1991) presented subjects with part-list cues in an order intended to either agree or conflict with the subjects' presumed subjective organization of list items. Only the conflicting order of part-list cues significantly inhibited recall relative to a free recall condition, suggesting that part-list cues do indeed affect the nature of the search set.

Nonrandom output order. The effects of an item's serial input position can only be partly reconciled with the search-set sampling account. The effects of input position effect on recall probability (primacy and recency effects) are not problematic; it may simply be the case that the earlier and later study items are more often included within the search set. But the effects of input position upon output position contrast with the assumption of equal retrieval probabilities. To account for this effect, the model would require additional assumptions—for example, multiple copies of early and late items within the search set. The simple pure-death process without extensions obviously cannot account directly for the effects of subjective organization or the interrelations between items held in long-term store as might a more complex model such as SAM (Raaijmakers & Shiffrin, 1980) or TODAM (Murdock, 1982). Nevertheless, as Vorberg and Ulrich (1987, pp. 5–6) observe, "when considered as a limiting case that applies only under conditions which minimize these factors, the model is remarkable in its ability to predict certain temporal properties of the retrieval from long-term memory." Indeed, the quality of fit provided by the ex-Gaussian distribution and the pure-death hyperbola in the present experiments is more than adequate. But more importantly, the utility of the pure-death process lies not in its ability to describe but rather in its ability to interpret. As Vorberg and Ulrich (1987, p. 3) conclude, "under carefully contrived experimental conditions that minimize the role of [long-term memory] organizational factors, the model may provide important insights into the dynamics of retrieval from long-term memory."

REFERENCES

ALBERT, D. (1968). Freies Reproduzieren von Wortreihen als stochastische Entleerung eines Speichers [Free recall of word sequences as stochastic depletion of storage]. *Zeitschrift für experimentelle und angewandte Psychologie*, **15**, 564-581.

ASHBY, F. G. (1982). Testing the assumptions of exponential, additive reaction time models. *Memory & Cognition*, **10**, 125-134.

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). San Francisco: W. H. Freeman.

BOUSFIELD, W. A., & SEDGEWICK, C. H. (1944). An analysis of restricted associative responses. *Journal of General Psychology*, **30**, 149-165.

BOUSFIELD, W. A., SEDGEWICK, C. H., & COHEN, B. W. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology*, **67**, 111-118.

BURBECK, S. L., & LUCE, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, **32**, 117-133.

DAWSON, M. R. W. (1988). Fitting the ex-Gaussian equation to reaction time distributions. *Behavior Research Methods, Instruments, & Computers*, **20**, 54-57.

GORFEIN, D. S., & JACOBSON, D. E. (1973). Memory search in a Brown-Peterson short-term memory paradigm. *Journal of Experimental Psychology*, **99**, 82-87.

GRONLUND, S. D., & SHIFFRIN, R. M. (1986). Retrieval strategies in re-

call of natural categories and categorized lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 550-561.

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.

HERRMANN, D. J., & PEARLE, P. M. (1981). The proper role of clusters in mathematical models of continuous recall. *Journal of Mathematical Psychology*, **24**, 139-162.

HOCKLEY, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 497-512.

HOCKLEY, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 598-615.

HOEL, P. G. (1971). *Introduction to mathematical statistics* (4th ed.). New York: Wiley.

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 sequences from long-term memory. *Psychological Review*, **77**, 317-331.

LUCE, R. D. (1986). *Response times: Their role in inferring elementary 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.

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.

METCALFE, J., & MURDOCK, B. B. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning & Verbal Behavior*, **20**, 161-189.

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

MURDOCK, B. B., JR., & OKADA, R. (1970). Interresponse times in single-trial free recall. *Journal of Experimental Psychology*, **86**, 263-267.

PAPOULIS, A. (1965). *Probability, random variables, and stochastic processes* (2nd ed.). New York: McGraw-Hill.

PATTERSON, K. E., MELTZER, R. H., & MANDLER, G. (1971). Interresponse times in categorized free recall. *Journal of Verbal Learning & Verbal Behavior*, **10**, 417-426.

POLLIO, H. R., KASSCHAU, R. A., & DENISE, H. E. (1968). Associative structure and the temporal characteristics of free recall. *Journal of Experimental Psychology*, **76**, 190-197.

POLLIO, H. R., RICHARDS, S., & LUCAS, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning & Verbal Behavior*, **8**, 529-536.

RAAIJMAKERS, J. G., & SHIFFRIN, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, pp. 207-262). New York: Academic Press.

RATCLIFF, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.

RATCLIFF, R. (1979). Group reaction distributions and an analysis of distribution statistics. *Psychological Bulletin*, **86**, 446-461.

RATCLIFF, R., & MURDOCK, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, **83**, 190-214.

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 & Memory*, **3**, 164-188.

ROEDIGER, H. L., III, & THORPE, L. A. (1978). The role of recall time in producing hypermnesia. *Memory & Cognition*, **6**, 296-305.

ROEDIGER, H. L., III, & TULVING, E. (1979). Exclusion of learned material from recall as a postretrieval operation. *Journal of Verbal Learning & Verbal Behavior*, **18**, 601-615.

SCHULZ, U., & ALBERT, D. (1976). Das Reproduzieren und seine Beenigung: Eine Klasse von Zählermodellen [Free recall and its interpretation: A class of counting models]. *Zeitschrift für experimentelle und angewandte Psychologie*, **23**, 678-699.

SHIFFRIN, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of human memory* (pp. 375-447). New York: Academic Press.

SLOMAN, S. A., BOWER, G. H., & ROHRER, D. (1991). Congruency effects in part-list cuing inhibition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 974-982.

TOWNSEND, J. T. (1971). A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, **10**, 161-163.

VORBERG, D., & ULRICH, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology*, **31**, 1-23.

WIXTED, J. T., & ROHRER, D. (1993). Proactive interference and the dynamics of retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1024-1039.

WIXTED, J. T., & ROHRER, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review*, **1**, 89-106.

APPENDIX A

The ex-Gaussian distribution, though conceptually quite simple, is rather unwieldy in form, which perhaps explains why it has appeared in several articles with typographical errors. Likewise, an outline of its derivation given in Luce (1986, p. 36), the only one that we have seen, includes a few typographical errors as well. For this reason, we present a full derivation below.

The ex-Gaussian distribution is derived by mathematically convolving the Gaussian distribution and the exponential distribution (cf. McGill, 1963, p. 315; Papoulis, 1965, pp. 134–136). In particular, if the random variables x, distributed as $f_x(x)$ for all x, and y, distributed as $f_y(y)$ for $y \geq 0$, then the random variable $t = x+y$ has the distribution

$$f_t(t) = \int_{-\infty}^{t} f_x(x) * f_y(t - x)dx.$$

If

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

and

$$f_y(y) = \lambda e^{-\lambda y} \text{ where } \lambda = \frac{1}{\tau},$$

then

$$f_t(t) = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \lambda e^{-\lambda(t-x)}dx$$

$$= \frac{\lambda}{\sigma\sqrt{2\pi}} * \int_{-\infty}^{t} e^{-\frac{(x^2-2x\mu+\mu^2)}{2\sigma^2}} e^{-\lambda\tau}e^{\lambda x}dx$$

$$= \frac{\lambda}{\sigma\sqrt{2\pi}} * e^{-\lambda t} \int_{-\infty}^{t} e^{-\frac{x^2-2x\mu+\mu^2-2\sigma^2\lambda x}{2\sigma^2}} dx$$

$$= \frac{\lambda}{\sigma\sqrt{2\pi}} * e^{-\lambda t}$$

$$\int_{-\infty}^{t} e^{-\frac{x^2-2x\mu+\mu^2-2\sigma^2\lambda x+2\mu\sigma^2\lambda+\sigma^4\lambda^2-2\mu\sigma^2\lambda-\sigma^4\lambda^2}{2\sigma^2}} dx$$

$$= \frac{\lambda}{\sigma\sqrt{2\pi}} * e^{-\lambda t} \int_{-\infty}^{t} e^{-\frac{x^2 - 2x\mu + \mu^2 - 2\sigma^2\lambda x + 2\mu\sigma^2\lambda + \sigma^4\lambda^2}{2\sigma^2} + \mu\lambda + \frac{\sigma^2\lambda^2}{2}} dx$$

$$= \frac{\lambda}{\sigma\sqrt{2\pi}} * e^{-\lambda(t-\mu) + \frac{\sigma^2\lambda^2}{2}} \int_{-\infty}^{t} e^{-\frac{[(x-\mu-\sigma^2\lambda)/\sigma]^2}{2}} dx$$

$$= \frac{\lambda}{\sqrt{2\pi}} * e^{-\lambda(t-\mu) + \frac{\sigma^2\lambda^2}{2}} \int_{-\infty}^{(t-\mu)/\sigma - \sigma\lambda} e^{-\frac{y^2}{2}} dy$$

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

or

$$\frac{1}{\tau\sqrt{2\pi}} * e^{-\frac{t-\mu}{\tau} + \frac{\sigma^2}{2\tau^2}} \Phi\left(\frac{t-\mu}{\sigma} - \frac{\sigma}{\tau}\right)$$

where $\Phi(x)$ is the Gaussian cumulative density function.

APPENDIX B

This appendix includes the derivation of exponential cumulative recall and exponentially declining rates of recall, given a random sampling model (cf. Bousfield & Sedgewick, 1944; McGill, 1963).

Let N equal the number of recoverable target items within the search set and let S equal the total number of items within the search set. Thus, the probability of choosing a recoverable target on the first sample is simply the ratio N/S. If the first sample was unsuccessful, the probability of success on the second sampling would remain N/S. If the first sample was successful, however, the probability of success on the second sampling would decline to $(N-1)/S$. If we let R equal cumulative recall, or the number of previously sampled recoverable targets, then the probability of success for any given sample equals

$$(N-R)/S. \tag{1}$$

This probability of success divided by the time needed to sample an item (t^*) yields the momentary probability of retrieval,

$$(N-R)/(St^*).$$

Replacing the constant St^* with the constant τ yields

$$(N-R)/\tau. \tag{2}$$

The momentary probability of retrieval, when interpreted as a continuous process rather than a discrete one, translates to the rate of recall,

$$r = (N-R)/\tau. \tag{3}$$

Since the rate of recall equals the derivative of cumulative recall, r can be replaced by dR/dt. Rearrangement yields the differential equation,

$$dR/(N-R) = (1/\tau)(dt).$$

Integrating both sides yields

$$-\ln(N-R) = (1/\tau)t + C,$$

where C is the constant of integration. After exponentiating both sides, rearrangement yields

$$R(t) = N - e^{-C}e^{-t/\tau}.$$

Since $R=0$ when $t=0$, e^{-C} must equal N. More rearrangement yields the cumulative exponential,

$$R(t) = N(1 - e^{-t/\tau}). \tag{4}$$

Differentiating Equation 4 and replacing dR/dt with r yields

$$r(t) = (N/\tau)e^{-t/\tau}. \tag{5}$$

Thus, random sampling yields exponential cumulative recall (4) and exponentially declining rates of recall (5). Furthermore, the mean latency of recall, given by the parameter in Equations 4 and 5, is seen to equal τ. Thus,

$$\text{mean latency or } \tau = St^*. \tag{6}$$

APPENDIX C

The nature of pure-death IRTs is discussed in detail in both McGill (1963) and Vorberg and Ulrich (1987). A brief derivation of the explicit equation that describes the growth of mean IRTs is given here.

As shown in Appendix B (Equation 2), the momentary probability of retrieval equals

$$(N-R)/\tau.$$

Thus, the momentary probability of retrieval decreases as the number of items recalled (R) increases. Once each item is recalled, the expected duration until the next item is recalled (i.e., the mean IRT) equals the reciprocal of the momentary probability of retrieval. Thus, the mean IRT that *follows* the Rth response equals

$$\tau/(N-R) \text{ for } R = 1, 2, \ldots N-1.$$

If we define the IRT that follows the Rth response as the ith, then the

$$\text{mean } i\text{th IRT} = \tau/(N-i).$$