

## Attended and Non-Attended States in Working Memory: Accessing Categorized Structures

Brian McElree

*New York University*

Following several recent approaches, working memory is argued to consist of a subset of representations that are within the scope of active processing and a larger set of recently activated, but non-attended representations. A response-signal speed-accuracy tradeoff (SAT) and complementary reaction time variant of a probe recognition task were used to measure the retrieval of categorized material from working memory. In SAT, two distinct retrieval speeds were found: A fast rate for items from the last studied category and a slower rate common to all items from other categories in the list. This difference in retrieval speed was evident in the shapes of the reaction time distributions. The faster retrieval speed for items from the last category is argued to reflect a matching process in which the recognition probe is rapidly compared to the current contents of awareness. © 1998 Academic Press

The relationship between awareness and working memory (WM) has been an issue since James (1890) equated the latter with the former: “*But an object from primary memory . . . was never cut off in consciousness from that of the immediate present moment. In fact it comes to us as belonging to the rearward portion of the present space of time . . .*” (pp. 643–647). In recent theorizing, however, a distinction is often drawn between WM structures that are the focus of active processing and those that have residual activation as a consequence of recent processing (e.g., Anderson, 1983; Baddeley, 1986; Conway & Engle, 1994; Engle, 1995; Cowan, 1988, 1995; Ericsson & Pennington, 1993; McClelland & Rumelhart, 1986; Schneider & Detweiler, 1988). The former are viewed as those structures that are the current object(s) of attention or awareness. The latter are structures that are outside the scope of attention, yet, as a consequence of recent processing, have a privileged

status over less recently processed long-term memory structures, either by having a temporary representation in a specialized store (e.g., Baddeley, 1986) or by having residual activation in a long-term representation (Anderson, 1983; Conway & Engle, 1994; Cowan, 1988, 1995; Engle, 1995).

The distinction between attended and non-attended WM representations has been indirectly motivated by the analysis of several tasks involving short-term retention (see Cowan, 1995). For example, Wickens, Moody, and Dow (1981), Wickens, Moody, and Vidulich (1985), and Halford, Maybery, and Bain (1988) have presented evidence that proactive interference is non-existent in small set sizes (less than 4 items) under conditions in which all items can plausibly be maintained in the current focus of attention (i.e., an immediate recognition test with no interpolated distractor task). Tehan and Humphreys (1995; see also, Tehan & Humphreys, 1996) demonstrated that proactive interference and phonological similarity are inversely related. On an immediate test, phonological similarity impacted on performance, but no proactive interference effect was found. Conversely, with a delayed test, proactive interference was found, but phonological similarity

The author thanks William Keleman and Rose Corona for assistance in running Experiment 1.

Address correspondence and reprint requests to Brian McElree, Department of Psychology, New York University, 6 Washington Place, 8th Floor, NY, NY, 10003. E-mail: bdm@psych.nyu.edu.

did not affect performance. Wickens et al. (1981, 1985) suggest that the lack of a proactive interference effect for small sets sizes tested immediately suggest that no retrieval processes are needed for items maintained in active awareness.

More direct evidence for the distinction between attended and non-attended states can be found in analyses of the speed with which various recently processed events can be accessed. Structures that are the current focus of processing should be more accessible than structures no longer active in awareness. In principle, this distinction should be testable with response time (RT) measures: Active structures should be associated with faster recognition latencies than structures that are no longer active in awareness (Cowan, 1995). Indeed, Wickens et al. (1981, 1985) reported that when a distractor task is interpolated between study and the recognition probe recognition latencies are longer for small set sizes (2 and 4 items) than when the recognition probe immediately follows study. Unfortunately, RT measures do not uniquely assess memory accessibility since RT can vary with item availability (strength, familiarity, fragility, etc.) (e.g., Doshier, 1979, 1982; McElree & Doshier, 1989; 1993; Wickelgren, 1977). It is possible that differences observed by Wickens et al. (1981, 1985) simply reflect underlying differences in memory availability.

The response-signal speed-accuracy trade-off (SAT) procedure provides a means of separately accessing retrieval speed (accessibility) and memory strength (availability). Several investigations of retrieval speed with SAT procedures have found systematic differences in accessibility that can be plausibly related to whether an item is within the current focus of awareness (Doshier, 1982; McElree, 1996; McElree & Doshier, 1989, 1993; Wickelgren, Corbett, & Doshier, 1980).

### *Retrieval Speed*

The response-signal SAT procedure measures retrieval speed and accuracy by cueing subjects to respond at various times after the

onset of a recognition probe. Subjects are required to respond within 300 ms of the response cue (typically a tone), whether or not processing is complete. Accuracy is thereby measured as a function of retrieval time. By sampling an appropriate range of times (e.g., 100–3000 ms), the full time-course of retrieval can be measured, including when accuracy departs from chance, the rate at which accuracy grows over retrieval time, and the ultimate or asymptotic level of performance. The asymptote of the retrieval function provides a measure of retrieval accuracy, the overall probability of recognition. The point at which accuracy departs from chance (intercept) and the rate at which accuracy grows provide joint measures of the speed or dynamics of retrieval. It is the latter (rate and/or intercept) that one would expect to vary between items that are or are not within the current focus of processing.

Wickelgren, Corbett, and Doshier (1980) used an SAT variant of a probe recognition task (e.g., Sternberg, 1966, 1975) to examine retrieval speed and accuracy for various serial positions within lists of 16 sequentially presented consonants. Asymptotic accuracy systematically decreased with less recent serial positions, consistent with standard forgetting models (see McElree & Doshier, 1989). Crucially, however, retrieval speed (SAT intercept and rate) was constant across all serial positions save the last, most recently studied position. The retrieval dynamics for the last study item was found to be 50% faster than the retrieval dynamics for every other serial position. Wickelgren et al. (1980) argued that, since no activity intervened between study and test, the most recent serial position remained active in awareness. Since the last item was active in awareness, the test probe could be directly compared to the current contents of awareness, circumventing the retrieval processes that were needed to restore items in a passive state to active processing.

The pattern of results reported by Wickelgren et al. (1980) were replicated and extended by McElree & Doshier (1989). McElree

& Doshier systematically examined the retrieval dynamics for all serial positions within list lengths of 3 to 6 words. Asymptotic accuracy decreased as the recognition probe was drawn from less recent serial positions, supplemented by a small primacy effect for the first item on the list. As with the Wickelgren et al. study, only two retrieval speeds were found: One for the last item on the list, and another for every other serial position (independent of list length). The retrieval dynamics for the most recently studied item were 44–55% faster than other serial positions. The fast dynamics for the most recent item support the Wickelgren et al. suggestion that the last item studied remains active in awareness. Crucially, McElree & Doshier (1989, Experiment 3) found that varying letter (upper versus lower) case between study and test did not attenuate the retrieval advantage for the last studied item. The lack of an effect of case suggests that the retrieval advantage was not mediated by a low level physical match (e.g., Posner, Boies, Eichelman, & Taylor, 1969), but rather by a more abstract representation of the studied item.

More compelling evidence that the retrieval advantage for the last studied item is not due to a low level physical match comes from a recent SAT comparison of recognition based on phonological and semantic properties. McElree (1996) presented subjects with 5-word lists, immediately followed by a recognition probe. Subjects were to judge whether the probe i) was in the memory set (item judgment), ii) rhymed with an item in the memory set (rhyme judgment), or iii) was a synonym of an item in the memory set (synonym judgment). For each type of judgment, the SAT retrieval functions showed a pattern that was similar to the data reported in McElree & Doshier (1989). Asymptotic accuracy decreased with decreasing recency of study, coupled with a small primacy effect. Uniform dynamics were observed for all serial positions except for the most recently studied position, which showed, as with prior studies, a large retrieval advantage. Comparisons across judg-

ments showed that synonym and rhyme judgments were associated with slower SAT dynamics than item judgments for all serial positions except the last, most recently studied position. In that case, equal dynamics were observed across the three judgments. This finding suggests that subjects were able to directly compare the recognition probe to phonological and semantic properties active in awareness.

### *Span of Active Processing*

The studies outlined above indicate that the last item studied in a sequentially presented list remains active in awareness and, as a result, is readily accessible to cognitive operations. This finding supports the idea that the architecture of WM enables a small subset of items to be the focus of active processing and awareness. However, does this imply that the focus of attention is restricted to one item or object?

Doshier (1981) reported a retrieval dynamics advantage for the last pair of items in a word-word paired associate recognition task. Minimally, this result indicates that at least two items can be active in awareness when the task demands require concurrent processing of two items. While clearly there may be an upper bound on the span of active awareness, the Doshier study suggests that what is active in awareness is the last unit of processing or chunk (Miller, 1956), which is determined by task requirements.

The studies reported here examine whether what remains active in awareness is contingent on the inherent structure of the studied material and how that structure interacts with memory encoding operations. An SAT (Experiment 1) and a complementary response time (Experiment 2) variant of a probe recognition task were used to examine retrieval of item information from lists with a categorical structure. Subjects were presented nine-item lists for study, consisting of three instances of each of three categories. The words in a list were sequentially presented, blocked by category membership. Following the presentation

of the list, subjects were presented a recognition probe and judged whether it was a member of the study set. The primary question addressed was whether the retrieval advantage observed in prior studies using a similar paradigm (McElree & Doshier, 1989; Wickelgren et al., 1980) would extend to the last category studied rather than the last item studied. If the retrieval advantage truly reflects the current contents of active processing, then all three items from the last category should remain active in awareness to the degree to which subjects use the categorical structure to encode items within WM.

### *Retrieving Categorized Structures*

A second, perhaps more basic question addressed in the current study is the nature of the mechanism that mediates retrieval from categorized structures. The McElree & Doshier (1989) analysis of the probe recognition task suggests that item recognition is mediated by a parallel or direct-access mechanism, such as a simple strength accumulator (Reed, 1973) or a parallel retrieval mechanism with a self-terminating decision rule (Ratcliff, 1978). The data reported in McElree & Doshier (1989) are incompatible with several models of short-term memory retrieval proposed on the basis of RT data, including serial exhaustive (Sternberg, 1975) or serial self-terminating (Theios, 1973) retrieval mechanisms, as well as rate-varying parallel mechanisms (Murdock, 1971, Townsend & Ashby, 1982). The crucial evidence that motivates the claim that retrieval is mediated by a parallel or direct-access retrieval mechanism is the finding that, except for the most recently studied item, retrieval dynamics (speed) does not vary for different serial positions or different list lengths. Serial exhaustive "scanning" mechanisms predict that retrieval speed should vary with set size. Serial self-terminating models and rate-varying parallel mechanisms predict that retrieval speed should vary with serial position or recency. McElree & Doshier (1989) found that serial position affected asymptotic accuracy only, consistent with the notion that the

strength of an item's representation in memory is affected by how recently it has been studied. These differences in item strength as a function of serial position average to produce set size differences in asymptotic accuracy. What Sternberg and others took to be differences in retrieval speed based on RT measures were shown to reflect differences in memory strength only.

A direct-access or parallel retrieval mechanism is consistent with several memory models that treat item recognition as an assessment of global familiarity, strength, or goodness-of-match (Gillund & Shiffrin, 1984; Hintzman, 1984, 1988; Murdock, 1982, 1983). However, several recent studies have documented boundary conditions on the applicability of parallel or direct-access retrieval mechanisms. Hintzman & Curran (1994) report data that indicate that a recall operation may supplement assessments of familiarity when discrimination is difficult (e.g., when plurals of a singular study item are used as lures). McElree, Jacoby, & Dolan (1996) found that a recollective (recall-based) operation is used to attenuate familiarity assessments when the recognition judgment necessitates source discriminations (e.g., determining whether an item was heard rather than read). Gronlund and Ratcliff (1989), Gronlund, Edwards, & Ohrt (1993), and Clark (1992) have argued that associative recognition is mediated by a recall operation performed on the component items of an associative probe. Finally, McElree & Doshier (1993) demonstrated that judgments of recency (Hacker, 1980, Muter, 1979) are based on a relatively slow serial self-terminating search process, rather than on a direct assessment of familiarity or strength.

Collectively, these studies suggest that the recovery of some types of information necessitate retrieval operations beyond a direct assessment of familiarity. Recovery of component elements within a categorized structure may represent one of these cases. For example, in the context of TODAM, a distributed memory model, Murdock (1993) argues that chunking at study forms a composite represen-

tation that is opaque in that the representation must be unpacked to recover component elements. Doshier & Rosedale (1989) examined priming effects with composite representations (word triples) and found evidence that indicated that component elements (a word) can prime the composite representation but not individual elements within the composite (i.e., individual words).

If component elements of a composite representation are directly accessible, one would expect a pattern of retrieval dynamics similar to what was reported by Wickelgren et al. (1980), McElree & Doshier (1989), and McElree (1996); specifically, uniform dynamics for all serial positions save those positions that are the focus of active processing. If elements are not directly accessible, then several patterns are possible. For example, composite representations could be serially scanned to identify a category that is compatible with the semantic properties of the test probe. Under these circumstances, retrieval dynamics will vary with the order in which the composite representations are scanned. Additionally or alternatively, a composite representation may need to be unpacked and component elements serially scanned. In this case, retrieval dynamics should vary for the elements within a composite category. Systematic examination of serial position profiles within a list allows one to test a number of these possibilities.

## EXPERIMENT I: SAT MEASURES

In the first experiment, the response-signal SAT procedure was used to examine retrieval speed and accuracy for each serial position within nine-word categorized lists. Figure 1 illustrates a sample trial. Following sequential presentation of the list and a short pattern mask, a test probe was presented for a “yes/no” recognition judgment. At one of seven lags (100–3000 ms after the onset of the test), a tone sounded and subjects were required to respond. Positive (“yes”) trials consisted of

test probes from one of the nine serial positions.

Five types of negative test probes were used to explore further properties of the retrieval mechanism that underlies the recognition of elements in a categorized structure. McElree and Doshier (1989), extending RT results reported by Atkinson, Hermann, and Westcourt (1974) and Monsell (1978), found that recently studied negative test probes (items drawn from the previous study list) induced elevated false alarm rates early in retrieval (100–900 ms) as compared to negative probes that were less recently studied (items drawn from lists three or more trials back). The higher false alarm rate early in retrieval was attenuated as more list-specific information accrued over the course of retrieval (times greater than 900 ms). These results indicate that the retrieval mechanism underlying item recognition is responsive to item strength (familiarity). Moreover, it indicates that in a short-term memory task retrieval processes are not restricted to the experimenter-defined memory set — a finding that is patently inconsistent with scanning mechanisms that are restricted to the current memory set (e.g., Sternberg, 1975; Theios, 1973). Here, a similar manipulation of recency was employed. One type of lure—a recent negative (RN)—was constructed by selecting words that had occurred on the previous study list. These lures were not a member of any of the three categories used on the current study list. Performance on RNs was compared to performance on distance negatives (DN), which were words from a category that was not used on the current study list nor the study list of the previous trial.

The RN-DN contrast examines the impact of episodically-based item strength or familiarity on recognition judgments. However, if the categorical structure of the lists is used to encode items, the retrieval mechanism may also be responsive to specific semantic properties that are used to encode the list. To examine whether semantic similarity impacts on the retrieval mechanism, unstudied members of

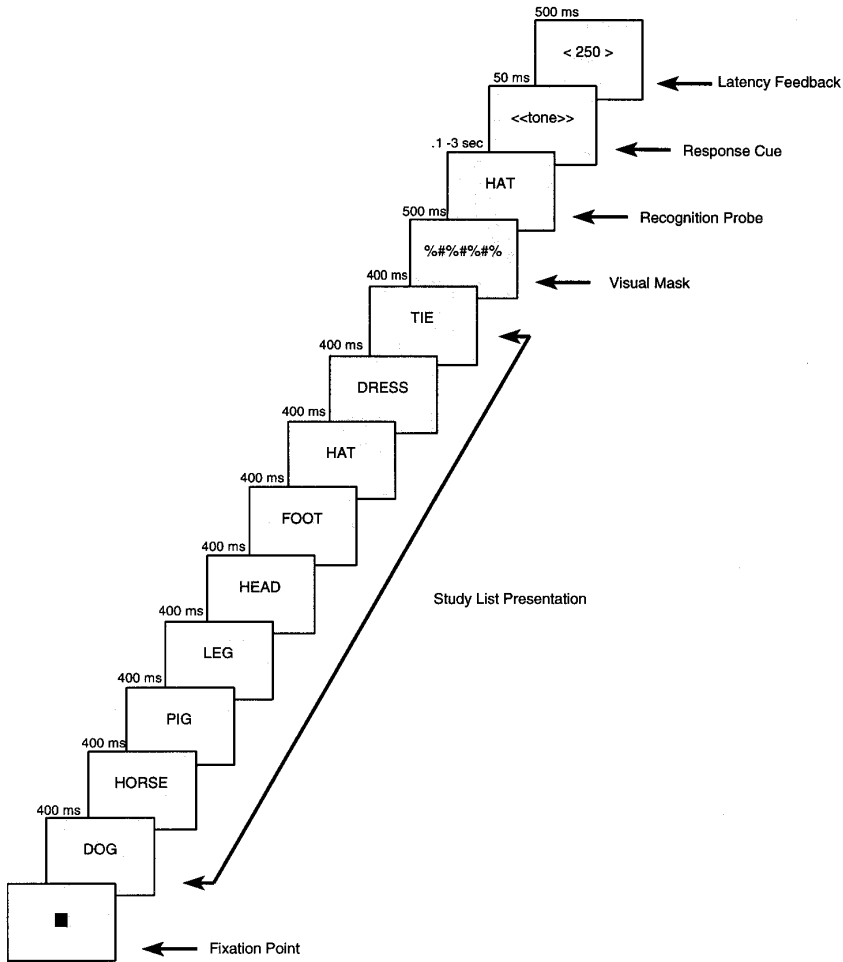


FIG. 1. A sample trial sequence illustrating the response-signal speed-accuracy tradeoff variant of the probe recognition task with categorized lists.

the first, second, and third categories on the list (denoted as 1N, 2N, 3N, respectively) were used as lures. These lures were compared to DNIs to assess whether the semantic similarity of a lure to a studied items/category induces a high false alarm rate comparable to the false alarm rate induced by episodic familiarity.

#### Method

*Subjects.* Six subjects from the University of California, Irvine and New York University served as subjects. Each subject participated

in 10 one-hour sessions, plus an additional 1-hour practice session that served as training for the SAT procedure. All subjects had normal or corrected vision.

*Materials.* Twelve one-syllable nouns from nine categories in the Battig and Montague (1969) category norms were used as stimuli. The Appendix lists the materials used in the experiments.

*Design and procedure.* Stimulus presentation and response collection were controlled by a personal computer. Each trial consisted of the presentation of a 9-word study list. A

study list was constructed by randomly sampling (without replacement) 3 words from 3 pseudo-randomly selected categories. The selection of categories excluded those that were used on the prior trial. An equal number of positive and negative probes was used. Positive probes were drawn from each of the 9 serial positions equally often. One-third of the negative trials used distant negatives (DN), lures drawn from categories not presented in the study phase nor in the study phase of the previous trial. Another third of the negative trials used recent negatives (RN), lures drawn from one of the 3 words used on the last category of the prior trial. The remaining third of the negative trials used words drawn from one of the 9 members of the categories on the study list that were not used as study items. Each of the three categories was sampled equally often (denoted 1N, 2N, 3N).

The sequence and timing of events within a trial is schematically represented in Figure 1 and were as follows: (1) A centered, square fixation point was presented for 500 ms. (2) Study words were sequentially presented for 400 ms in the center of an otherwise blank screen. (3) The final word in the study list was masked by a collection of nonletter symbols presented for 500 ms. (4) The test word was presented in the same region as the study list and mask. (5) The test word remained on the screen for either 43, 200, 300, 500, 800, 1500, or 3000 ms, at which point the screen cleared and a 50-ms (2000 Hz) tone sounded to cue the subjects to respond. Subjects responded by pressing one of two keys on a numeric keypad to denote either that the test item had appeared in the study list ('3' key) or that it was not in the study list ('1' key). (6) Following a response, visual feedback on the subject's latency to respond to the interruption tone was presented. Subjects were instructed to respond within 270 ms of the tone. They were informed that responses longer than 270 ms were too long and that responses faster than 120 ms were anticipations. Subjects initiated the next trial by pressing a key.

Each of the 10 approximately one-hour ses-

sions consisted of 720 trials, divided into 4 blocks of 180 trials. For each subject, this yielded 40 trials at each of the seven lags for each of the 9 positive conditions, as well as for the 1N, 2N, and 3N negative trials. Both the DN and RN lure conditions had 120 trials for each of the seven interruption lags. This number of trials provided sufficient data to analyze individual subject data.

### *Results and Discussion*

For each subject, the hit rate for each serial position at each response lag was scaled against the respective false alarm rate for the distant negatives (DN) to derive an (equal variance-Gaussian)  $d'$  measure. Perfect performance at any lag was adjusted by a minimum-error correction that replaced a zero error rate with an error rate of 0.05 as a means of ensuring that, given the sample size, the  $d'$  values were measurable (Macmillan & Creelman, 1991). An alternative  $d'$  scaling is discussed below (Discriminative Analysis).

*Asymptotic accuracy.* Performance at the two longest interruption lags (1.5 and 3 s) provides an empirical measure of asymptotic recognition accuracy. Figure 2 shows the average (over subjects) asymptotic  $d'$  (solid line and filled squares) for each of the nine serial positions, along with the data from each of the six subjects (dashed line and open symbols). An analysis of variance on the individual subjects' asymptotic  $d'$ s indicated that performance significantly declined ( $\alpha = .05$ ) as the test probe was drawn from less recently studied serial positions [ $F(8,40) = 42.69$ ,  $MSe = 0.0801$ ]. In addition to this general trend toward poorer performance with decreasing recency, there is a tendency for bowed serial position functions within each category. To test the significance of this latter tendency, the data were fit with a regression model. A simple linear fit using serial position as the single predictor of  $d'$  performance accounted for 91.3% of the variance. (Plausible transformations of the serial position predictor, e.g., a logarithmic transformation, reduced the quality of fit.) Adding an additional predictor that

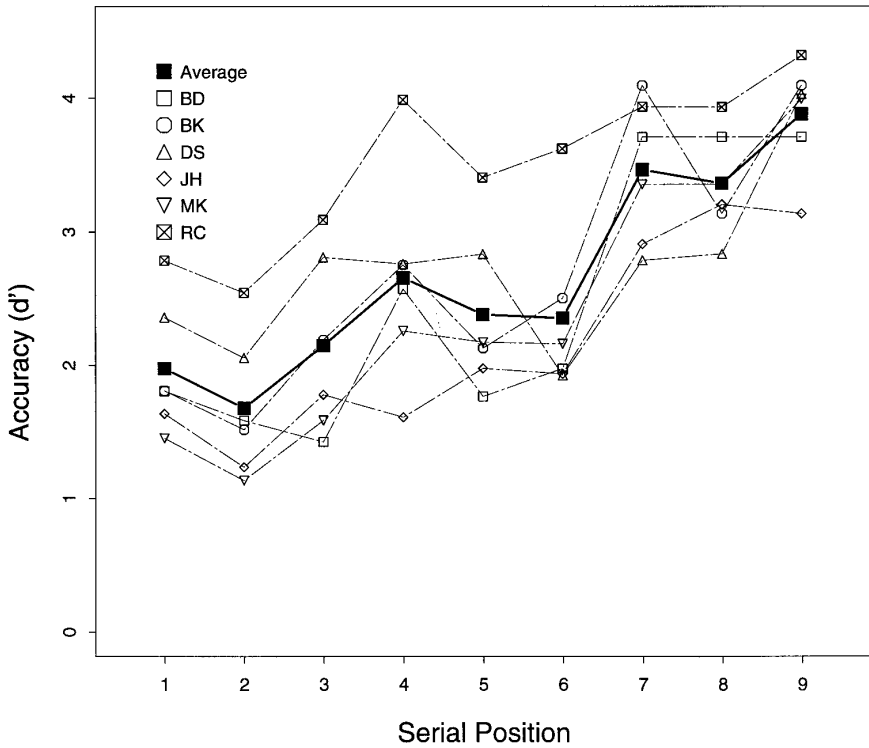


FIG. 2. Observed average asymptotic  $d'$  values as a function of the serial position of the positive test probe.

coded the first serial position of a new category to mark the transition from one category to another (i.e., serial positions 1, 4, and 7 were coded as 1, while all other positions were coded as 0) substantially improved the fits, accounting for 96.3% of the variance. The partial correlation of the additional predictor was significant,  $t(6) = 2.77$ ,  $p = .0321$ , indicating that the asymptotic profiles are partly a function of the categorical structure of the list.

McElree and Doshier (1989) found that a simple strength model originally proposed by Wickelgren and Norman (1966) for untimed forgetting functions provided an accurate fit of asymptotic performance as a function of recency. The *acquisition-primacy* model assumes that item strength for a particular serial position ( $k$ ) is:

$$d(k, L) = \alpha(k)\phi^{L-k} \quad (1)$$

where  $L$  is list length,  $k$  is serial position,  $\alpha(k)$  is a serial-position dependent acquisition parameter, and  $\phi$  is a decay parameter that is a function of the number of items intervening between study and test.<sup>1</sup> The model assumes that two factors control item strength: A serial position-dependent acquisition parameter and a per-item decay constant. McElree and Doshier (1989) found a decay rate ( $\phi$ ) of 0.82 in fits of the serial position profiles from list lengths of 3-6 words. The acquisition parameter was highest for the first item on the list, and decreased with each successive position. The counteracting influences of decay and the acquisition parameters produced bowed serial position functions, showing steady decrease

<sup>1</sup> Wickelgren (1970) demonstrated that forgetting also depends on rate of presentation in addition to the number of intervening items.

in asymptotic accuracy coupled with a small primacy effect.

The design of this experiment, having only one list length, is too unconstrained to provide a strong test of the applicability of the acquisition-primacy model. However, it is noteworthy that when the average data are fit with this model, the estimated decay rate ( $\phi$ ) is .84, which is quite similar to the .82 rate found in McElree and Doshier (1989). The primary difference here is that the acquisition parameters ( $\alpha(k)s$ ) do not systematically decline with additional serial positions:  $\alpha(1) = 4.5$ ,  $\alpha(2) = 3.3$ ,  $\alpha(3) = 3.5$ ,  $\alpha(4) = 3.7$ ,  $\alpha(5) = 2.8$ ,  $\alpha(6) = 2.4$ ,  $\alpha(7) = 2.9$ ,  $\alpha(8) = 2.4$ ,  $\alpha(9) = 1.3$  in  $d'$  units. While there is an overall decline, there are reversals when a new category is introduced (i.e., serial position 4 and serial position 7). These reversals, of course, serve to capture the within-category bowing of the serial position functions in Figure 2 and correspond to the additional predictor in the regression analysis above. The model's treatment of these effects indicates that shifting to a new category increases the acquisition strength of new category members. In turn, this could suggest that a new category draws more attentional resources or that a shift of category induces a within list release from proactive interference.

*Retrieval dynamics.* Figure 3 shows the average (over subjects) empirical  $d'$  values (symbols) as a function of retrieval time (the lag of the interruption tone plus the latency to respond to the tone). Inspection of Figure 3 suggests that performance on the last three serial positions, the last category on the list, are markedly different from the remaining serial positions. The question that is addressed here is whether this difference rests solely in asymptotic accuracy or whether the items from the last category are associated with faster recognition dynamics, either SAT rate, SAT intercept, or some combination of the two.

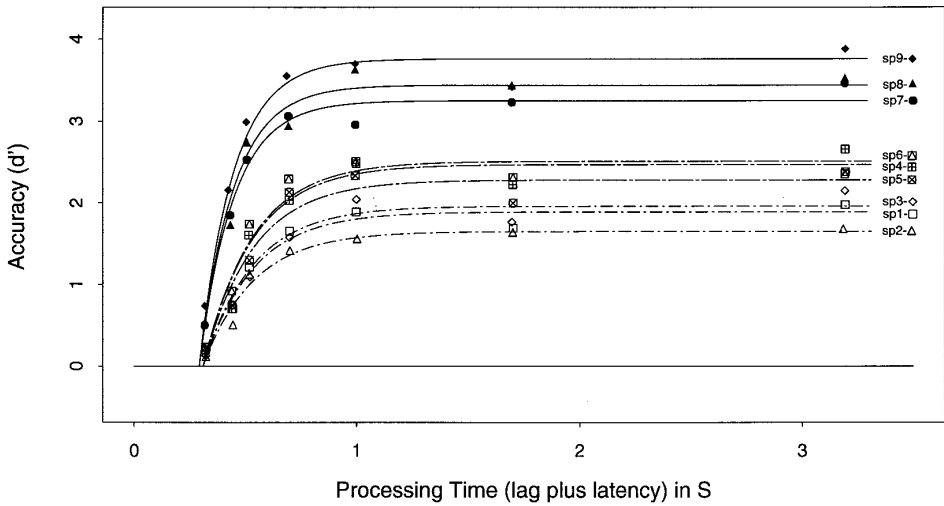
To estimate retrieval dynamics (speed) for each serial position, the full time-course SAT

functions were fit with an exponential approach to a limit:

$$d'(t) = \lambda(1 - e^{-\beta(t-\delta)}), \text{ for } t > \delta, \text{ else } 0. \quad (2)$$

Equation 2 describes the growth of accuracy over retrieval time using three parameters: i)  $\lambda$  an asymptotic parameter reflecting the overall probability of recognition; ii)  $\delta$  an intercept parameter reflecting the discrete point in time when accuracy departs from chance ( $d' = 0$ ); and iii)  $\beta$  a rate of rise parameter that describes the rate at which accuracy grows from chance to asymptote. Differences in retrieval speed or dynamics are reflected in either the intercept ( $\delta$ ) and/or rates of rise to asymptote ( $\beta$ ) parameters. The intercept and rate of an SAT function reflect either the rate of continuous information accrual or the distribution of finishing times of a discrete or quantal process (Doshier, 1976, 1979, 1981, 1982, 1984; Meyer, Irwin, Osman, & Kounois, 1988; Ratcliff, 1988). Numerous studies have found that Equation 2 provides a precise quantitative summary of the shape of a full time-course SAT function (Doshier, 1976, 1979, 1981, 1982, 1984; McElree, 1993, 1996; McElree & Doshier, 1989, 1993; McElree & Griffith, 1995; Reed, 1973, 1976; Wickelgren, 1977; see also Ratcliff, 1978 for an alternative three-parameter equation derived from the random-walk (diffusion) model and McElree & Doshier, 1989 for a comparison of the two equations).

All analyses were performed on individual subject data. Consistent patterns across subjects are summarized with analyses and graphs of the average (over subjects) data. Differences among the serial position SAT functions were quantified by fitting the exponential in Equation 2 with an iterative hill climbing algorithm (Reed, 1976), similar to STEPIT (Chandler, 1969). This fitting procedure minimized the squared deviations of predicted values from observed data. A hierarchical model testing scheme was used to determine the best fitting exponential model. The nine retrieval



**FIG. 3.** Average  $d'$  accuracy (symbols) as a function of processing time (lag of the response cue plus latency to respond to the cue) for each serial position. Smooth curves show the best fits of Equation 2 with the (average) parameters listed in Table 1.

functions were fit with sets of nested models that systematically varied the three parameters of Equation 2. These models ranged from a null model in which all functions were fit with a single asymptote ( $\lambda$ ), rate ( $\beta$ ), and intercept ( $\delta$ ) to a fully saturated (27-parameter) model in which each function was fit with a unique asymptote, rate, and intercept. The quality of the fit was assessed by using three criteria, as follows: i) The value of an  $R^2$  statistic,

$$R^2 = 1 - \frac{\sum_{i=1}^n (d_i - \hat{d}_i)^2 / (n - k)}{\sum_{i=1}^n (d_i - \bar{d})^2 / (n - 1)}, \quad (3)$$

where  $d_i$  represents the observed data values,  $\hat{d}_i$  indicates the predicted values,  $\bar{d}$  is the mean,  $n$  is the number of data points, and  $k$  is the number of free parameters (Reed, 1973). This  $R^2$  statistic is the proportion of variance accounted for by the fit, adjusted by the number of free ( $k$ ) parameters (Judd & McClelland, 1989). ii) Evaluation of the consistency of the parameter estimates across the subjects. iii) Evaluation of whether the fit yielded systematic (residual) deviations that could be accommodated by allocating more (i.e., separate) parameters to various conditions.

Given the impact of serial position on asymptotic accuracy, it is not surprising that all models that did not allocate separate asymptotic parameters ( $\lambda$ s) to each serial position left systematic residuals and produced extremely poor fits to the empirical SAT data. All viable models needed minimally to allocate a separate  $\lambda$  to each serial position. In the average data, allocating separate  $\lambda$ s to each serial position increased the adjusted- $R^2$  value from .576 to .953. Comparable differences were seen in the fits of all subjects. The crucial issue concerned the allocation of the dynamics parameters, rate ( $\beta$ ) and intercept ( $\delta$ ).

For the average data, a  $9\lambda-2\beta-2\delta$  exponential model gave the best representative fit of the empirical data. This model allocated a separate asymptote ( $\lambda$ ) to each serial position, one rate ( $\beta$ ) and intercept ( $\delta$ ) to positions 7, 8, and 9—the most recently studied category on the list—and another rate and intercept to the remaining positions (1–6). This model yielded an adjusted- $R^2$  value of .973, which is substantially better than a  $9\lambda-1\beta-1\delta$  model (.953) that assumed uniformed dynamics across the list. This improvement in  $R^2$  resulted from the additional dynamics parameters capturing high

performance at early interruption times for serial positions 7–9, performance that was systematically underestimated by the  $9\lambda-1\beta-1\delta$  fit. Additionally, the  $9\lambda-2\beta-2\delta$  model yielded higher adjusted- $R^2$  values than any other allocation of dynamics parameters, including higher parameter variants of the exponential extending up to a fully saturated  $9\lambda-9\beta-9\delta$  model. Perhaps most importantly,  $9\lambda-2\beta-1\delta$  model that allocated a separate rate parameter to the last serial position (rather than category)—the model that produced the best fit to the data in McElree & Doshier (1989) and McElree (1996)—or a related  $9\lambda-1\beta-2\delta$  that varied intercept rather than rate both yielded lower adjusted- $R^2$  values (.961 and .961, respectively) than the  $9\lambda-2\beta-2\delta$  model. Crucially, these models also underestimated performance at early interruption times for serial positions 7 and 8.

In the  $9\lambda-2\beta-2\delta$  model, the average intercept for items from the last category was estimated at 295 ms as compared to 312 ms for items from the two prior categories (serial positions 1–6). The average rate parameter for items from the last category was estimated at 6.47 as compared to 4.45 for items from the two prior categories. When the two dynamics estimates are combined into a composite measure of retrieval speed ( $\delta + 1/\beta$ ), the retrieval dynamics for items from the last category was estimated at 449 ms as compared to 537 ms for items from the two less recent categories. The smooth functions in Figure 3 show the best fitting  $9\lambda-2\beta-2\delta$  exponential model, using the (average) parameters listed in Table 1.

The  $9\lambda-2\beta-2\delta$  fit of the average data reflects the fact that, with one exception discussed below, fits of all individual subjects' data showed clear evidence that items from the last category were associated with relatively fast judgment dynamics. However, the  $9\lambda-2\beta-2\delta$  fit of the average data is a generalization across individual subjects and does not directly reflect the pattern seen in any particular subject. Three subjects (BK, JH, RC) were best fit by a model that simply allocated a

separate intercept to items from the last category, *viz.*  $9\lambda-1\beta-2\delta$  model (adjusted- $R^2$ s of .928, .919, .907, respectively). Two subjects (BD, MY) were better fit by a model that allocated a separate rate to items from the last category, *viz.*  $9\lambda-2\beta-1\delta$  model (adjusted- $R^2$ s of .907 and .949, respectively). Whether dynamics differences are expressed in intercept or rate can have important theoretical implications in some circumstances. McElree & Doshier (1993), for example, argued that intercept differences motivate a serial model for the retrieval of order information. However, in the present application little hinges on whether the “true” difference lies in rate or intercept, in that in both cases the data indicates that items from the last category are associated with faster responses than other serial positions.<sup>2</sup> This rate/intercept ambiguity has been found in other studies of uncategorized materials. For example, Wickelgren et al. (1980) found that the advantage for the last item could be equally captured in rate or intercept. McElree & Doshier (1989) found marginally better fits when rate was varied, while McElree (1996) found marginally better fits when intercept was varied.

Irrespective of whether the advantage in dynamics were best captured in rate or intercept, allocating a separate dynamics parameter for items from the last category substantially improved the overall quality of fit in all five subjects relative to other possible allocations of the dynamics parameters. The estimated advantage for the last category across the five subjects ranged from 41–172 ms using  $\delta + 1/\beta$  as a measure of composite retrieval speed. Crucially, in contrast to studies with uncategorized material (Wickelgren et al., 1980; McElree & Doshier,

<sup>2</sup> Whether a difference in retrieval speed engenders differences in intercept or rate can depend on several factors including, for example, the size of the variance relative to the size of the mean differences. If the variance is large relative to the mean difference, the advantage is likely to be expressed in rate. Conversely, if the variance is small relative to the size of the difference, differences are likely to be expressed in intercept.

TABLE 1  
Exponential Parameter Estimates

Parameter	Average	BD	BK	DS	JH	MY	RC
Common scaling							
Serial position functions							
Serial Position 1 $\lambda$	1.89	2.09	1.62	2.39	1.29	1.73	2.27
Serial Position 2 $\lambda$	1.66	1.78	1.24	1.91	1.31	1.42	2.32
Serial Position 3 $\lambda$	1.96	1.59	1.92	2.38	1.58	1.71	2.71
Serial Position 4 $\lambda$	2.47	2.57	2.29	2.57	1.77	2.05	3.55
Serial Position 5 $\lambda$	2.29	2.24	1.89	2.29	1.93	2.16	3.18
Serial Position 6 $\lambda$	2.51	2.34	2.51	2.23	2.10	2.26	3.59
Serial Position 7 $\lambda$	3.25	3.66	3.35	2.81	2.87	2.95	3.91
Serial Position 8 $\lambda$	3.44	3.64	3.46	2.93	3.08	3.36	4.34
Serial Position 9 $\lambda$	3.76	3.99	3.73	3.07	3.38	4.22	4.34
Serial Position 1–6 $\delta + 1/\beta$	0.537	0.505	0.515	0.669	0.528	0.557	0.589
Serial Position 7–9 $\delta + 1/\beta$	0.449	0.373	0.474	0.669	0.429	0.451	0.451
Adjusted- $R^2$	.973	.907	.928	.886	.919	.947	.907
Intrusion functions							
1N Lure $\lambda_1$	0.11	0.52	0.5	-0.95	-0.12	-0.39	0.69
2N Lure $\lambda_1$	0.54	2.01	0.66	0.05	0.09	0.17	0.83
3N Lure $\lambda_1$	0.95	2.80	0.89	2.08	1.04	0.64	0.82
RN Lure $\lambda_1$	0.53	1.42	0.52	0.267	0.37	0.5	0.41
Common $\lambda_2$	0.17	0.14	0.01	0.29	0.12	-0.2	0.44
Common $\delta_2$	0.556	0.396	0.803	0.338	0.624	1.004	0.510
Discriminative scaling							
Serial position functions							
Serial Position 1 $\lambda$	1.77	1.81	1.58	2.05	1.39	2.21	1.69
Serial Position 2 $\lambda$	1.52	1.48	1.16	1.57	1.39	1.89	1.73
Serial Position 3 $\lambda$	1.84	1.25	1.95	2.05	1.68	2.19	2.14
Serial Position 4 $\lambda$	2.17	2.32	2.15	2.25	1.56	2.04	2.92
Serial Position 5 $\lambda$	1.99	1.93	1.71	1.97	1.72	2.15	2.56
Serial Position 6 $\lambda$	2.21	2.05	2.38	1.91	1.89	2.25	2.95
Serial Position 7 $\lambda$	3.11	3.77	3.06	2.60	2.58	2.82	3.66
Serial Position 8 $\lambda$	3.33	3.75	3.19	2.72	2.78	3.27	4.12
Serial Position 9 $\lambda$	3.65	4.07	3.47	2.88	3.09	4.17	4.12
Serial Position 1–6 $\delta + 1/\beta$	0.583	0.591	0.690	0.702	0.523	0.581	0.483
Serial Position 7–9 $\delta + 1/\beta$	0.551	0.506	0.576	0.702	0.501	0.536	0.461
Adjusted- $R^2$	.968	.889	.904	.814	.861	.775	.851

1989; McElree, 1996), the  $R^2$  values for the respective models for each of the five subjects were higher than alternative models in which a separate dynamics parameter (either intercept or rate) was allocated to the last item [respectively:  $R^2 = .907$  versus .882 and .881 for BD;  $R^2 = .927$  versus .917 and .914 for BK;  $R^2 = .919$  versus .889 and .900 for JH;  $R^2 = .947$  versus .942 and .943 for MY;  $R^2 = .907$  versus .889 and .896 for RC]. The addition of more parameters beyond one dynamics parameter for items from the last

category did not yield any consistent pattern across subjects and, as with the average data, reduced the  $R^2$  value in all cases. Hence, there was no evidence to indicate that there were any dynamics differences among positions 1-6 or among positions 7-9.

The data from one (DS) of the six subjects showed a discrepant pattern. There was no clear evidence to support the notion of a retrieval advantage for items from the last category. A  $9\lambda-1\beta-2\delta$  model, in which a separate intercept was allocated to items from the last

category, produced slightly higher adjusted- $R^2$  values than a simple  $9\lambda-1\beta-1\delta$  model (.886 versus .885). However, an equivalent  $R^2$  was produced by an alternative  $9\lambda-2\beta-1\delta$  model in which a separate rate was allocated to the last item rather than category. Given the near equal quality of the different fits, it is not possible to clearly discriminate among possible models for this subject.

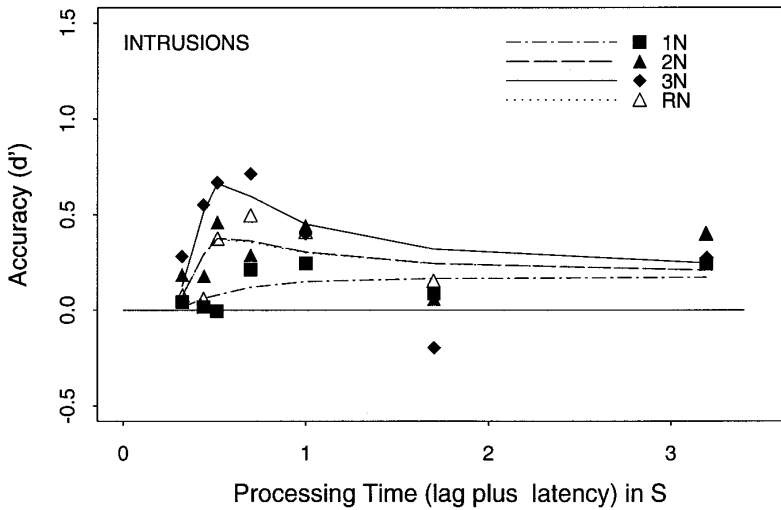
Given that one of the six subjects did not show clear evidence of a dynamics advantage for items from the last category, a direct difference t-test was used to test the significance of the estimated difference between the dynamics for serial positions 7–9 and 1–6. The combined dynamics estimate ( $\delta + 1/\beta$ ) was used to place rate and intercepts in an equal metric. [Since neither a  $9\lambda-1\beta-2\delta$  model or a  $9\lambda-2\beta-1\delta$  model could be motivated from the analysis of subject DS's data, estimates from the  $9\lambda-1\beta-1\delta$  model were used in the t-test. Hence, for this subject the same value was used for position 1–6 and 7–9.] The t-test yielded a significant difference,  $t(5) = 13.06$ ,  $p < .05$ , indicating that items from the last category were indeed associated with faster recognition judgments than all other positions.

Finally, it should be noted that while the adjusted- $R^2$  value for the  $9\lambda-2\beta-2\delta$  model differs from a model without dynamics differences (*viz.*,  $9\lambda-1\beta-1\delta$  model) by a substantial amount (.953 versus .973), the differences in adjusted- $R^2$  among models with alternative allocations of the dynamics parameters are substantially smaller. In the average data, for example, the difference between the  $9\lambda-2\beta-2\delta$  model shown in Figure 3 and models that allocate one rate or intercept to the last item rather than category is .973 versus .961 and .961, respectively. Note, however, that these models are simultaneously fitting 63 data points (9 serial positions crossed with 7 interruption lags). A systematic misfit of the intercept (or rate) for two of the three items from the last category will underestimate 2–4 of these 63 data points at best, and this will not have a dramatic effect on  $R^2$ , particularly an adjusted- $R^2$ . To illustrate why the small im-

provement in adjusted- $R^2$  is in fact a meaningful one, the three functions from the last category were fit separately, reducing the data points from 63 to 21. When the rate and intercept were set to the same values as the other 6 serial positions (estimated from the  $9\lambda-1\beta-1\delta$  model of all 63 data points), the adjusted- $R^2$  value was .949. When the estimates were taken from a  $9\lambda-1\beta-2\delta$  or a  $9\lambda-2\beta-1\delta$  model in which a separate dynamics parameter was allocated to the last item rather than category, the adjusted- $R^2$  value was similarly .949. In contrast, the adjusted- $R^2$  value increased to .968 when the estimates from the  $9\lambda-2\beta-2\delta$  model were used. The markedly higher adjusted- $R^2$  value illustrates that the model is indeed capturing substantial, systematic differences associated with the last three serial positions.

Selectively fitting the last three positions also provides a more sensitive test of whether there are residual differences among positions 7–9 that were not detected in the composite fits. A series of competitive fits ranging from a  $1\lambda-1\beta-1\delta$  model to a fully saturated  $3\lambda-3\beta-3\delta$  model were performed in which each of the parameters was free to vary. In five of the six subjects, a  $3\lambda-1\beta-1\delta$  model gave the highest adjusted- $R^2$  value, indicating that there was no evidence for any systematic differences among positions 7–9 other than the asymptotic differences evident in Figure 2. One subject (MY) was better fit by a  $3\lambda-3\beta-1\delta$  model (.944 versus .935). MY's estimated rates for serial position 7–9 were 5.8, 4.3, 6.1, respectively. However, if we ignore differences in  $R^2$  and examine the parameter estimates for comparable models across the other subjects, this ordering is only seen in one other subject and no other consistent trend emerged. Therefore, there is no compelling evidence to motivate a claim that the dynamics vary within the last category.

*Episodic and semantic intrusions.* To examine whether the retrieval process is responsive to episodically-based item strength and semantic similarity, the z-scores for false alarm rate for a recent negative (RN,



**FIG. 4.** Average pseudo- $d'$  values (symbols) for recent negatives (RN) and lures from the first (1N), second (2N), and third (3N) categories on the list. Smooth curves show the best fits of Equation 4 with the (average) parameters listed in Table 1.

an item from the prior list) and unstudied members of the three categories on the lists (1N, 2N, and 3N) were scaled against the  $z$ -score for the false alarm rate for a distance negative (DN, a lure that was not a member of the categories on the list nor words presented on the prior list). The scaling produces what Doshier, McElree, Hood, & Rosedale (1989) and McElree & Doshier (1989) termed a pseudo- $d'$  measure, where higher  $d'$  values denote poorer performance resulting from the intrusion of strength or similarity into the decision process.

The symbols in Figure 4 show the pseudo- $d'$  scaling for the average (over subjects) data. With the possible exception of the 1N-DN condition, all functions tended to be nonmonotonic in that they show an early rise to a high false alarm rate (200-800 ms), followed by a period later in retrieval when the false alarm

rate is attenuated (pseudo- $d'$  approaches zero). These nonmonotonic false alarm rates indicate that the retrieval mechanism is sensitive to item strength and semantic similarity early in retrieval. The correction later in retrieval indicates that the intrusion of strength or similarity is corrected by the accrual of list-specific information.

Nonmonotonic SAT functions motivate a two-process retrieval model. To quantify the differences in intrusion across the various conditions, a dual-process variant of the exponential model was applied to the data. Ratcliff (1980) derived a two-process SAT model from the diffusion model (see Doshier et al. 1989 for an application to a priming paradigm), and in McElree & Doshier (1989) this approach was adapted to the exponential form:

$$d'(t) = \begin{cases} \lambda_1(1 - e^{-\beta(t-\delta_1)}), & \text{for } \delta_1 < t < \delta_2 \\ \lambda_2 + \frac{(\lambda_1 - \lambda_2)(\delta_2 - \delta_1)}{(t - \delta_1)} \times (1 - e^{-\beta(t-\delta_1)}), & \text{for } t \geq \delta_2. \end{cases} \quad (4)$$

Equation (4) states that during an initial time slice ( $\delta_1 < t < \delta_2$ ) response accuracy is controlled by the accrual of one type of information, in this application the accrual of either episodic or semantic familiarity. Response accuracy during this period is modelled as a simple exponential approach to an asymptote ( $\lambda_1$ ). At time  $\delta_2$ , however, additional list-specific information begins to accrue. The net effect is to shift response accuracy from the asymptote  $\lambda_1$  operative during the first period ( $\delta_1 < t < \delta_2$ ) to a new asymptote  $\lambda_2$ . The second part of Equation 4 estimates this new asymptote and when the shift in processing begins ( $\delta_2$ ).<sup>3</sup>

To quantify the intrusions (false alarms) stemming from recent presentation (RN) and semantic similarity (1N, 2N, and 3N), Equation 4 was used to estimate  $\lambda_1$ ,  $\lambda_2$ , and  $\delta_2$  in the average and individual pseudo- $d'$  data. The dynamics parameters from the respective fits of the serial position functions were used for  $\delta_1$  and  $\beta$ . For each type of lure, a unique  $\lambda_1$  was estimated to assess the magnitude of the intrusion, while a common  $\lambda_2$  and  $\delta_2$  were used for all four conditions. To the degree that recent presentation or the semantic similarity of a lure induces a high false alarm rate relative to the baseline DN condition,  $\lambda_1$  and  $\lambda_2$  estimates will be above zero. If episodic strength and semantic similarity intrude more early in retrieval than late, then  $\lambda_1$  will be estimated to be higher than  $\lambda_2$ .

The smooth functions in Figure 4 show fits of Equation 4 to the average data using the parameters listed in Table 1. Recent study of a lure produced early intrusions estimated at  $\lambda_1 = 0.53 d'$  units (ranging from 0.27 to 1.42 across subjects). Semantic similarity also induced a high false alarm rate early in retrieval, with a magnitude that depended on how re-

cently the category was studied. The  $\lambda_1$  parameter was highest for members of the last category (3N), estimated 0.95 (ranging from 2.8 to 0.64), next highest for members of the next to last category (2N), estimated 0.53 (ranging from 0.05 to 2.01), and moderate to nonexistent for the first category studied (1N), estimated at 0.11 (ranging from -0.95 to 0.69). All subjects showed this ordering, save RC who showed a slightly higher  $\lambda_1$  for 2N as compared to the 3N lures (0.83 versus 0.82). These differences in intrusions rate were significant in an ANOVA performed on the three  $\lambda_1$  estimates,  $F(2,10) = 7.8$ ,  $MS_e = 0.345$ . The common  $\lambda_2$  was estimated to be 0.17 (ranging from -0.13 to 0.35), substantially lower than  $\lambda_1$  for all conditions save lures that were members of the first category studied. The higher  $\lambda_1$  values produce the nonmonotonic form of the functions shown in Figure 4. The parameter  $\delta_2$ , the point at which the asymptote shifts from  $\lambda_1$  to  $\lambda_2$ , was estimated at 556 ms (ranging from 338 to 1004 ms). In all but one subject (DS),  $\delta_2$  was estimated to be later than  $\delta_1$  [ $t(5) = 2.73$ ,  $p < .05$ ].

The intrusion analysis indicates that during an initial phase of recognition (200–800 ms), recognition processes are responsive to both episodic strength and semantic similarity. The intrusion of episodic strength has been documented in several SAT studies, including McElree & Doshier (1989) and McElree, Jacoby, & Dolan (1996). Intrusions based on semantic similarity (category membership) is a new finding. Intrusions based upon category membership provide independent evidence that subjects used the category structure to encode the elements on the list. The relatively high intrusion rate ( $\lambda_1$ ) for 3N and 2N lures demonstrates that recognition processes are at least as susceptible to semantic similarity as they are to episodic strength.

*Discriminative analysis.* The intrusion analysis indicates that the false alarm rate for a lure drawn from a category on the list varies directly with how recently the category was studied. Scaling the full retrieval functions against lures that are not members of any cate-

<sup>3</sup> Equation (4) makes two simplifying assumptions, namely equal rates ( $\beta$ ) for the two sets of processes and that processing shifts at a discrete point ( $\delta_2$ ). Both assumptions could be relaxed, but in practice there are simply not enough internal constraints in empirical data to motivate these higher parameter variants of (4).

gory on the list (DN) ignores these differences in false alarm rate. It could be argued that the retrieval advantage for the last category on the list reflects nothing more than a bias to respond positively to items that are semantically congruent with the last category studied. For example, Doshier et al. (1989) found that priming episodic recognition judgments with semantically related words produced a biased effect early in retrieval: The prime increased the hit rate for a semantically related target word but by an amount that was completely off-set by a corresponding increase in the false alarm rate for a semantically related lure. Scaling the hit rate for various serial positions against the false alarm rate for lures from the respective category provides a means of determining whether or the degree to which the retrieval advantage is due to a bias. Following Doshier et al. (1989), this scaling is referred to as a *discriminative* scaling.

Figure 5 shows the average (over subjects) asymptotic discriminative- $d'$  scalings for each of the nine serial positions (solid line and filled squares), along with the corresponding data from each of the six subjects (dashed line and open symbols). As before, an ANOVA on the individual subjects' asymptotic  $d'$  indicated that performance significantly declined as the test probe was drawn from less recently studied serial positions [ $F(8,40) = 47.69$ ,  $MSe = 0.0854$ ]. A simple linear fit using serial position as the single predictor of  $d'$  performance accounted for 91.6% of the variance. Adding an additional predictor that coded the first serial position of a new category improved the fits, accounting for 95.2% of the variance, but the partial correlation of the additional predictor failed to reach significance,  $t(6) = 2.06$ ,  $p = .084$ . However, the same trend toward within-category bowing of the recency profiles evident in Figure 2 is seen in Figure 5.

Figure 6 shows the full retrieval functions for the average data using the discriminative- $d'$  scaling. Analyses of the individual retrieval functions with competitive fits of Equation 2 indicated that the retrieval advantage for the last category is not due solely to a bias effect.

That is, as with the prior  $d'$  scaling, models which allocated a separate dynamics parameter (in this case a separate  $\delta$ ) to items from the last category yielded higher adjusted- $R^2$  values than did models that assumed no dynamics differences or models with alternative dynamics parameter allocations. This pattern was evident in five of the six subjects, with the exception being again subject DS, who did not show clear evidence of any dynamics differences.

The average data were best fit with a  $9\lambda-1\beta-2\delta$  model in which a separate intercept was allotted to items from the last category. The smooth functions in Figure 6 show the fits of this model using the parameters listed in Table 1. This model produced the highest adjusted- $R^2$  value, although the difference in  $R^2$  from a model which assumed no dynamics advantage (*viz.*,  $9\lambda-1\beta-1\delta$  model) was small (.968 versus .963; but see above). The estimated intercept for positions 7–9 was 293 ms as compared to a 325 ms intercept for positions 1–6. With the common  $d'$  scaling (DN used as the false alarm rate), the dynamics difference in the average data for the last as compared to other categories on the list was 88 ms. The reduction from 88 to 32 ms indicates that a sizable portion of the dynamics advantage for the last category was due to a bias to respond positively to items congruent with the last category studied.

The  $9\lambda-1\beta-2\delta$  model for the average data is motivated by the fact that this model provided the best fit across all subjects, save subject DS. Like the average data, subjects BD, JH, MY and RC showed a smaller dynamics advantage with the discriminative scaling [84 ms versus 132 ms, 22 ms versus 99 ms, 46 ms versus 106 ms, and 21 ms versus 172 ms, respectively]. Subject BK showed a larger advantage under the discriminative scaling (114 ms versus 41 ms). A t-test on the estimated intercepts was significant,  $t(5) = 2.63$ ,  $p = 0.046$ . [Again, a common intercept estimate from the  $9\lambda-1\beta-1\delta$  model was used for subject DS. If the  $9\lambda-1\beta-2\delta$  estimates were used,  $t(5) = 3.61$ ,  $p = .015$ .]

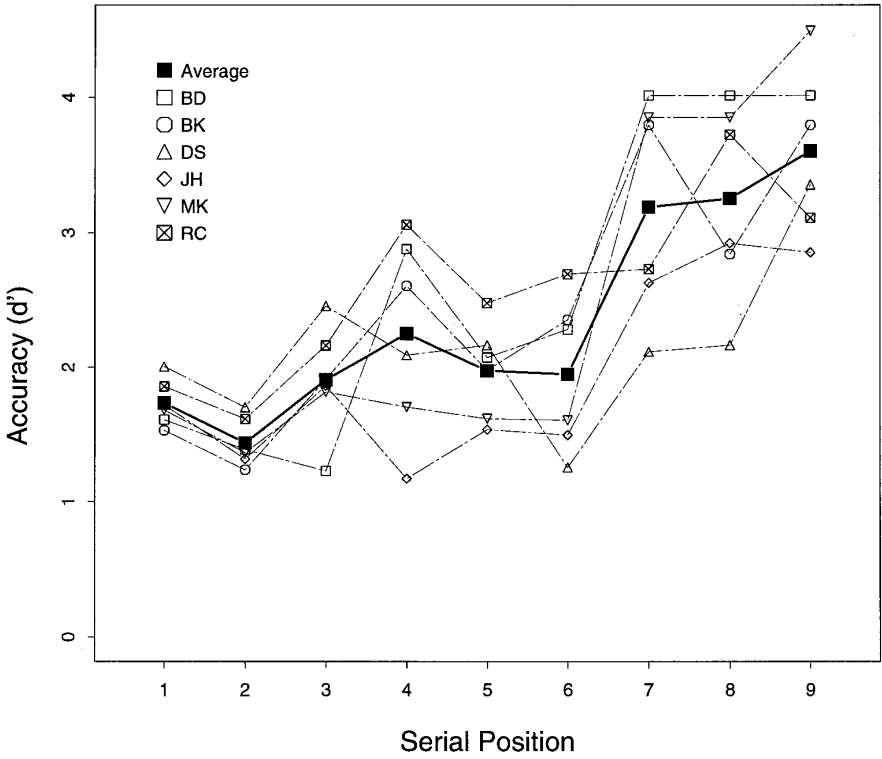


FIG. 5. Observed average asymptotic discriminative- $d'$  values as a function of the serial position of the positive test probe.

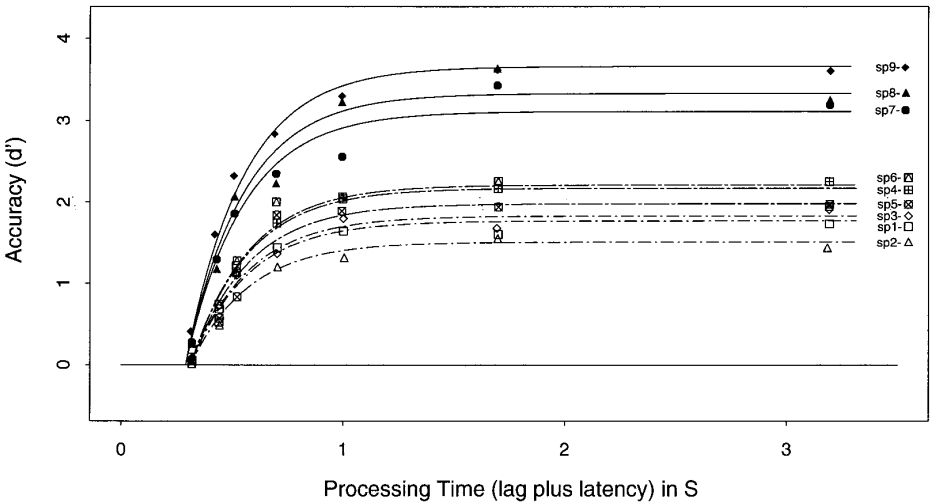


FIG. 6. Average discriminative- $d'$  accuracy (symbols) as a function of processing time (lag of the response cue plus latency to respond to the cue) for each serial position. Smooth curves show the best fits of Equation 2 with the (average) parameters listed in Table 1.

*SAT summary.* SAT asymptotes decreased as the (positive) test probe was drawn from less recently studied positions, consistent with standard forgetting models (e.g., McElree & Doshier, 1989; Wickelgren & Norman, 1966). Additionally, the asymptotes showed slight bowing within categories, suggesting that the internal structure of the list affected the strength of an item's representation in memory. With the exception of one subject, the SAT dynamics showed clear evidence of a retrieval advantage (SAT rate or intercept) for items from the last category on the list relative to the retrieval speed for other list positions. Beyond this dichotomy, there was no evidence to suggest that retrieval dynamics varied with the recency of categories or the items within a category. The retrieval advantage for items from the last category supports the notion that a few items can remain within the focus of attention.

## EXPERIMENT 2: RT MEASURES

Cowan (1995) and Wickens et al. (1981, 1985) have suggested that items that are the current focus of attention should be associated with faster response times than items outside the focus of attention. While this is likely true, RT measures also vary with the strength of an item's representation in memory. Doshier (1984) and McElree & Doshier (1989) found, for example, that mean RT varied with SAT asymptotic accuracy alone. Consequently, it is difficult to use differences in mean RT to motivate a distinction between attended and non-attended representations when items also differ in their underlying strength.

While differences in mean RT can not motivate such a distinction, differences among the higher order shape parameters of the RT distribution may provide additional evidence that is compatible with the strong evidence found in the analysis of SAT dynamics. Ratcliff & Murdock (1976; see also Luce 1986a) have argued that RT distributions are adequately described by an Ex-Gaussian function:

$$f(t) = \frac{e^{-\left[\frac{t-\mu}{\tau}\right] + \frac{\sigma^2}{\tau^2}}}{\tau\sqrt{(2\pi)}} \int_{-\infty}^{(t-\mu)/\sigma - \sigma/\tau} e^{-y^2/2} dy \quad (5)$$

Equation 5 represents the convolution of a Gaussian and an exponential distribution. The two parameters of the Gaussian, namely the mean  $\mu$  and the variance  $\sigma$ , describe the leading edge and mode of the RT distribution. The parameter  $\tau$  of the exponential describes the rightward tail or skew of the distribution.

Empirically, the parameters of Equation 5 have been found to roughly correspond to systematic changes in the shape of the SAT function. McElree & Doshier (1993) found that large SAT dynamics differences in a relative judgment of recency paradigm, in particular shifts in the SAT intercept ( $\delta$ ), produced concomitant shifts in the leading edge and mode of the respective RT distributions (reflected in changes in  $\mu$  and  $\sigma$  in Equation 5). Doshier (1984, and unpublished RT-distribution data) found that visually embedding test items in (masking) characters induced a slowing of the SAT rate parameter that was also mirrored by changes in the leading edge and mode of the RT distribution. In contrast, experimental factors that only affect the asymptote of the SAT function have been found to primarily impact on the tail of the RT distribution, reflected in changes in the  $\tau$  parameter. For example, Hockley (1984; Hockley & Corballis, 1982) found that set size in a probe recognition task did not affect the leading edge or mode of the distribution but only the  $\tau$  or skew of the distribution, with larger set sizes shifting the tail of the distributions toward longer times. This finding suggests that set size affects only a proportion of trials rather than increasing the duration of a serial component common to all trials. In SAT, set size affects asymptotic accuracy rather than inducing changes in retrieval dynamics (McElree and Doshier, 1989).

The correspondence between parameters of Equation 5 and the shape of the SAT function is less than perfect, for a number of principled reasons. The shape of an RT distribution may be strongly dependent upon the particular

speed-accuracy criterion used in the RT task (see Luce, 1986b and McElree & Doshier, 1993, and Ratcliff, 1978 for specific illustrations). Within the context of specific retrieval models, factors that control one of the SAT parameters may induce changes in all three parameters of the RT distribution. For example, in Ratcliff's (1978) diffusion model—one of the few explicit retrieval models that seeks to jointly account for RT distributional and SAT data—changes in the resonance parameter (the degree of match between the test probe and an element in memory) affects the asymptote of the SAT function. While this parameter largely impacts on the tail of the RT distribution, reflected in  $\tau$ , large differences in resonance will also induce small changes in the leading edge and mode of the distribution.

While the correspondence between the shape of the SAT function and the shape of the RT distribution is not perfect, it is nevertheless plausible to expect that conditions that show fast SAT dynamics will produce RT-distributions that have an earlier mode and leading edge. Ratcliff (1978) presented RT-distributions for serial positions with set sizes of 3–5 items in a probe recognition task. Although specific parameter values were not reported, visual inspection of the plots suggests that the distributions for the last serial position show an earlier mode and leading edge than other serial positions. Serial positions beyond the most recent differ from one another in terms of the skew of the distributions. The most recent serial position is, of course, the position that Wickelgren et al. (1980), McElree & Doshier (1989) and McElree (1993) found was associated with fast SAT dynamics. The current experiment examined whether the distributions for items from the last category differ from other items in terms of the mode and leading edge. If the last category is active in awareness such that test probes be can directly matched against the contents of awareness, then the RT-distributions for these items should show an earlier leading edge and mode than the distributions for items that require additional retrieval operations.

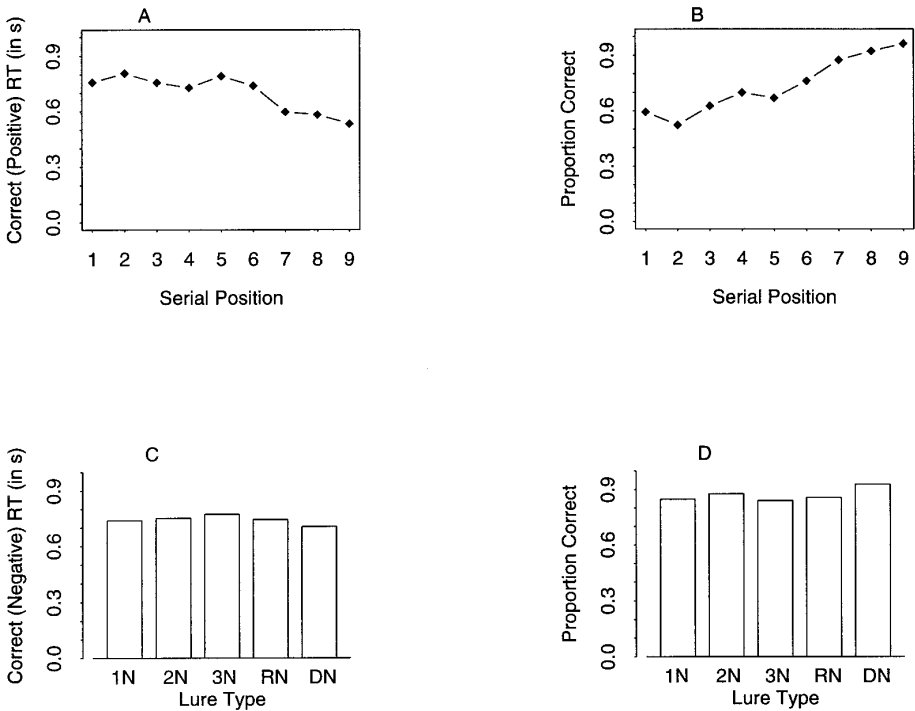
## Method

*Subjects.* Four subjects from New York University served as subjects. Each subject participated in 4–5 one-hour sessions. All subjects had normal or corrected vision.

*Materials, design, and procedure.* The material, design and procedure for this experiment were the same as Experiment 1 with the exception that the test probe was not followed by a response cue (tone) and remained on the screen until subjects responded. Subjects were instructed to respond to the test probe “as quickly but as accurately as they could”.

## Results and Discussion

*Mean RT and proportion correct.* Figure 7 shows the average latency to respond to the test probe and the proportion correct for positive trials (Panels A and B, respectively) and the five types of lures (Panels C and D, respectively). For positive trials the serial position of the test probe significantly ( $\alpha = 0.05$ ) impacted on mean RT [ $F(8,24) = 22.06$ ,  $MS_e = 0.0018$ ] and proportion correct [ $F(8,24) = 10.18$ ,  $MS_e = 0.0093$ ]. Mean RT increased and accuracy decreased as the test probe was drawn from less recently studied serial positions. As with the SAT asymptotic profiles, there was some bowing of the functions for items within the first and second category. A simple linear fit of the RT and proportion correct data using serial position as the single predictor accounted for 85.4% and 95.9% of the variance, respectively. Again, adding an additional predictor to code the first element of each category increased the proportion of variance accounted for (89.8% and 97.9% respectively). However, the partial correlation of the additional predictor was marginally significant in the proportion correct data [ $t(6) = 2.41$ ,  $p = .053$ ] but non-significant in the RT data [ $t(6) = 1.54$ , n.s.]. Although lures that were members of the categories on the list (1N, 2N, 3N), as well as recent negatives (RN), yielded slightly longer RTs and lower accuracy than distant negatives, these differences were not significant. However, given the



**FIG. 7.** Mean correct response time and proportion correct for positive test probes (Panel A and B, respectively) and negative test probes (Panel C and D, respectively) from the reaction time variant of the probe recognition task (Experiment 2).

low number of subjects, a test of means has little statistical power.

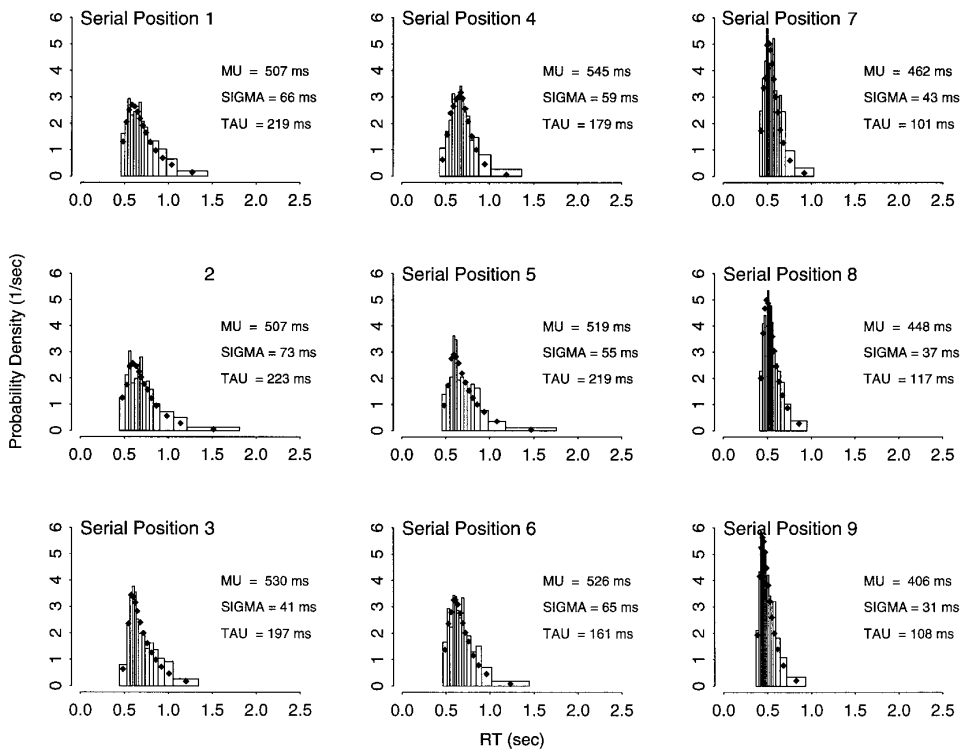
*RT distributions.* To examine how the shape of the RT-distribution varied as a function of the serial position of the positive probes, RT-distributions were constructed for each serial position by defining 15 equal-probability bins and adjusting the width of the respective bins. This procedure yields more stable estimates than does the alternative method of defining equal-intervals and calculating response frequency within each interval (see Ratcliff, 1979). The distributions were constructed for individual subjects, and a group (average) distribution was constructed by vincentized averaging (Ratcliff, 1979).

Figure 8 shows the group distributions (bars) for each of the 9 serial positions. Visual inspection of the distributions suggests that judgments of items from the last category were indeed associated with an earlier leading

edge and mode. That is, the distributions for serial positions 7–9 are shifted to left relative to the distributions for serial positions 1–6 and the former are more condensed than the latter.

To quantify these differences, Equation 5 was fit to each of the nine distributions for the group data and the individual subject's data. Table 2 lists the  $\mu$ ,  $\sigma$ , and  $\tau$  parameter estimates for the group and individual subject's data. The solid diamonds in Figure 8 show the predicted probability density for each serial position based on the parameter estimates for the group data.

The compression of the RT-distributions for serial positions 7–9, the last category, evident in Figure 8 was clearly reflected in the fits of the ex-Gaussian model to the group and individual subject data. Figure 9 plots the ex-Gaussian parameter estimates for the group data. With a few small reversals (e.g., serial position 3 in the



**FIG. 8.** The average (vincentized) reaction time (RT) distributions as a function of the serial position of the positive test probe. The bars show the estimated probability density at each of 15 time quantiles spanning the range of RTs. Solid diamonds show fits of the ex-Gaussian model (Equation 5) with the (average) parameters in Table 2.

group data), the parameter  $\sigma$ , the variance of the Gaussian component in Equation 5, increased as the test probe was drawn from less recently studied positions. The parameters  $\mu$ , the mean of the Gaussian, and  $\tau$ , the rate of the exponential, also show a trend to increase with less recent study positions. However, both parameters show a clear breakpoint between serial position 7 and serial position 6. In the case of  $\mu$ , although there are modest differences among the  $\mu$  parameter estimates for serial positions 7–9, these differences are small in comparison to the large differences between positions 7–9 and positions 1–6. Likewise, the  $\tau$  parameter is approximately constant for serial positions 7–9 and is much lower for positions 7–9 than for positions 1–6. These fits indicate that the RT-distributions for judgments of items from the last category on the list have an earlier mode and are less skewed

than the distributions for judgments of the items from the first categories.

The pattern seen in the group data is reflected in the data from all four subjects. Although there are some modest discrepancies (likely reflecting parameter tradeoffs), there is a clear breakpoint seen in the  $\mu$  and the  $\tau$  parameter estimates for serial positions 7–9 and positions 1–6 in all cases. The consistency of the pattern is impressive given that each serial position was independently fit, with no constraints applied across the fits.

## GENERAL DISCUSSION

### *Empirical Summary*

The response-signal SAT procedure provides measures of asymptotic accuracy, the overall probability of recognition, and re-

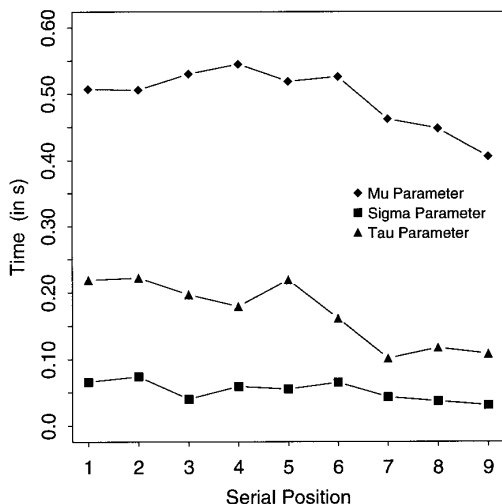
TABLE 2  
Ex-Gaussian Parameter Estimates

	Group	DA	JH	TY	WR
Serial position 1					
$\mu$	0.507	0.447	0.405	0.454	0.718
$\sigma$	0.066	0.082	0.031	0.044	0.163
$\tau$	0.219	0.178	0.208	0.200	0.230
Serial position 2					
$\mu$	0.507	0.459	0.445	0.439	0.792
$\sigma$	0.073	0.066	0.091	0.051	0.110
$\tau$	0.223	0.259	0.123	0.189	0.216
Serial position 3					
$\mu$	0.530	0.475	0.446	0.474	0.799
$\sigma$	0.041	0.051	0.013	0.052	0.039
$\tau$	0.197	0.242	0.156	0.191	0.245
Serial position 4					
$\mu$	0.545	0.483	0.429	0.483	0.874
$\sigma$	0.059	0.054	0.039	0.059	0.188
$\tau$	0.179	0.177	0.147	0.145	0.126
Serial position 5					
$\mu$	0.519	0.466	0.387	0.493	0.755
$\sigma$	0.055	0.062	0.036	0.049	0.069
$\tau$	0.219	0.239	0.236	0.169	0.219
Serial position 6					
$\mu$	0.526	0.485	0.401	0.473	0.739
$\sigma$	0.065	0.056	0.026	0.059	0.101
$\tau$	0.161	0.161	0.181	0.112	0.289
Serial position 7					
$\mu$	0.462	0.401	0.361	0.393	0.659
$\sigma$	0.043	0.041	0.014	0.035	0.071
$\tau$	0.101	0.101	0.095	0.101	0.293
Serial position 8					
$\mu$	0.448	0.394	0.346	0.394	0.685
$\sigma$	0.037	0.051	0.010	0.035	0.082
$\tau$	0.117	0.116	0.103	0.099	0.161
Serial position 9					
$\mu$	0.406	0.323	0.371	0.372	0.584
$\sigma$	0.031	0.018	0.035	0.013	0.070
$\tau$	0.108	0.090	0.090	0.124	0.168
Adjusted- $R^2$ <sup>a</sup>	.873	.845	.718	.827	.631

<sup>a</sup> Adjusted- $R^2$  value represents fits of all 9 serial position using a separate  $\mu$ ,  $\sigma$ , and  $\tau$  for each distribution (27 parameters for 126 data points).

trieval dynamics or speed. Experiment 1 demonstrated that both recognition speed and recognition accuracy are affected by recency. However, recency does not impact on recognition speed and accuracy in the same manner. Save for a small effect of the categorical structure of the list, recognition accuracy continuously declined with the time since study and/

or the number of items intervening between study and test. This general pattern is consistent with simple forgetting models (e.g., Wickelgren & Norman, 1966; McElree & Doshier, 1989) in which recency affects the strength of an item's representation in memory. The categorical structure of the list slightly perturbed this trend by producing



**FIG. 9.** Estimated  $\mu$ ,  $\sigma$ , and  $\tau$  parameters from fits of the ex-Gaussian model to the average (vincentized) reaction time distributions as a function of the serial position of the positive test probe.

small within-category bowed serial position profiles.

In contrast, recognition speed, estimated by SAT dynamics, did not continually decline with recency, but rather showed a dichotomous pattern in which items from the last category were associated with a fast retrieval rate and all items from other categories were associated with another, slower retrieval rate. This dichotomy has been demonstrated in other related SAT studies (Doshier, 1981; McElree & Doshier, 1989; McElree, 1996; Wickelgren et al., 1980). However, in prior cases the fast retrieval rate has been restricted to the last test item when no other items intervened between study and test. Here, the retrieval advantage was shown to extend beyond the last study item to encompass other items that were members of the same semantic category.

In the RT task, mean RT increased and proportion correct decreased when the test probe was drawn from less recently studied serial positions. Both measures appropriately mirror the pattern seen in the SAT asymptotes. The SAT asymptote reflects the strength of an item's representation in memory. The RT pat-

tern is consistent with classes of models in which strength differences engender differences in both mean RT and RT-accuracy (e.g., McElree & Doshier, 1989; Murdock & Dufty, 1972; Ratcliff, 1978). Differences in retrieval speed, estimated by SAT dynamics, will also impact on mean RT and perhaps RT-accuracy. For example, McElree & Doshier (1989) found that SAT asymptotes predicted mean RT for conditions that differ in asymptote only, but systematically overestimated mean RT for conditions that were also associated with a faster SAT rate.

Given that mean RT represents a mixture of strength-based effects and retrieval speed effects, the dichotomy in SAT between retrieval speed for the last category and all others is not directly observable with this measure alone. Analysis of the shape of the RT-distributions, however, showed that judgments concerning items from the last category yielded distributions that were tightly compressed around a mode that was substantially earlier than the mode for judgments of items from other categories. Hence, as with other comparison of SAT and RT-distributions (Doshier, 1984; McElree & Doshier, 1993), SAT dynamics differences were reflected in the leading edge and mode of the RT-distribution.

#### *Retrieving Categorized Structures*

Wickelgren et al. (1980) argued that a fast rate for the last item processed results from a matching process in which the encoded recognition probe is directly matched to the current contents of awareness. When an item is active in awareness, retrieval processes are not needed to restore the item to active processing. Items outside the current focus of attention require a retrieval operation to enter into active processing. The slower SAT dynamics associated with these items is interpretable as the time to retrieve an item from outside the focus of active processing. Residual activation resulting from recent processing of an item no longer within the current scope of attention plausibly controls the probability of restoring an item to active processing, and hence will

be reflected in the SAT asymptote. Residual activation does not, however, appear to affect the speed of retrieval.

For lists with categorical structures, the retrieval advantage extends to the last category processed rather than the last item. This extended retrieval advantage is likely a consequence of using the category structure to code items into a semantic chunk that in turn serve as the focus of active processing. Independent support for semantic coding based on category structure is evidenced by the finding of (albeit modest) effects of category structures on asymptotic accuracy and by the finding of relatively high false alarm rates early in retrieval to lures that were member of the categories studied on the list. Aside from the extended retrieval advantage, there was no indication that the retrieval of categorized structures induced a type of retrieval process distinct from the parallel or direct-access mechanism that has been argued to mediate retrieval from less structured material. SAT and RT measures for categorized lists show a pattern comparable to other analyses of item recognition (e.g., McElree & Doshier, 1989; McElree, 1996; Wickelgren et al., 1980). In particular, there was no evidence to indicate that items within a category or the categories themselves were scanned with a serial mechanism.

### *WM Architectures*

Several theorists (e.g., Anderson, 1983; Conway & Engle, 1994; Cowan, 1988, 1995; Engle, 1995) have suggested that WM consists of activated long-term memory structures, a subset of which form the current focus of attention. The finding of a retrieval advantage that extends to the last semantic category on a list and the associated RT-distributional patterns provide temporal data that support an architectural distinction between attended and non-attended memory representations. However, the data may also be consistent with other WM architectures that draw a similar distinction.

Baddeley and colleagues (Baddeley & Hitch, 1974; Baddeley, 1986) have proposed

an architecture that consists of a central executive—an attentional control system—that executes and coordinates processes that operate on two slave systems, namely a visuospatial sketch pad for the formation and maintenance of visuospatial images and a phonological loop for the formation and maintenance of verbal (speech-based) information. One extension of this approach to the current work would be to argue that those representations that are associated with fast retrieval speeds are just those that the central executive has selected for current processing. Indeed, Baddeley (1993) has recently argued that the central executive might be accurately described as “working attention”.

Consideration of the Baddeley model raises the issue of whether the retrieval advantage could be due to explicit rehearsal of items from the last category on the list. Baddeley (1986) argues that short-term verbal tasks, such as the probe recognition task, are mediated by a verbal component of WM that consists of a phonological store and a phonological rehearsal mechanism. The function of the latter is to maintain and refresh items in the fast-decaying phonological store. [Baddeley, Lewis & Vallar (1984) argued that auditorily presented digits decayed within 1 or 2 seconds.] Items that showed a retrieval advantage may be just those items that were within the rehearsal loop at test time. (See Seamon & Wright (1976) for an examination of the effects of rehearsal strategies on RT in a probe recognition task.)

A rehearsal-based explanation of the current results is not inconsistent with the claim that a subset of items remains within the current focus of attention, for rehearsal represents nothing other than active processing. However, rehearsal is likely only one type of cognitive operation that could effectively maintain items within the scope of attention. Moreover, it is questionable whether a rehearsal-based explanation is fully compatible with the studies reported here as well as related studies such as Wickelgren et al. (1980). First, the retention interval used in the current work was

brief (500 ms) and the relatively fast presentation rate (400 ms/item) may not have provided sufficient time to rehearse all three items from the last category. Minimally, it would appear necessary to argue that, across trials, subjects randomly select items from the last category to rehearse. Second, there is no ready explanation for why subjects would rehearse three items in the current study but not in related studies that used the same retention interval (McElree & Doshier, 1989; McElree, 1995) or a longer retention interval (700 ms in Wickelgren et al., 1980). The primary difference between the current study and prior studies is the categorized structure of the list. Consequently, it is more likely that specific encoding operations were primarily responsible for maintaining more than one item within the scope of attention.

It is also important to note that the Baddeley model (1986) proposes three possible states for an item; specifically, a representation in long-term memory, a representation in the phonological store, and/or a representation in the rehearsal loop. Although Baddeley and colleagues have not specifically addressed issues concerning retrieval time, it is natural to assume that the model would predict three distinct retrieval speeds associated with each of the three representational systems. This does not, however, appear to be the case. In the current work and in the related work of Wickelgren et al. (1980), McElree & Doshier (1989) and McElree (1995) only two retrieval speeds have been found for item recognition: One fast speed that is associated with a few items that are argued to be the focus of current processing and another, slower speed for all items outside the scope of attention. Apart from the fast dynamics for items within the scope of attention, retrieval speed remains constant for all other items in lists as large as 9 items in the current study and up to 16 items in the Wickelgren et al. (1980). Moreover, as McElree & Doshier (1989; see also Doshier & McElree, 1992) note, there is no evidence from the examination of retrieval speed and accuracy to suggest that the type of mecha-

nism that mediates recognition over the short-term is distinct from the type of mechanism that mediates recognition of long-term events. Given these facts, a parsimonious treatment of the SAT retrieval data is to assume a simple dichotomy such as the distinction between items that are the current focus of attention versus those items that are outside the span of attention.

Of course differences in retrieval speed do not necessarily force one to appeal to a special representational state such as the focus of attention. An alternative account might argue, for example, that the retrieval advantage for items from the last category reflect the fact that these items match well with current contextual cues; specifically, the cues established by processing of the last item on the list. In contrast, items from other categories would be incongruent with these contextual cues and may require alternative cues to enable successful retrieval. In this and related notions, one appeals to the contents of current awareness as the determinant of available retrieval cues rather than an alternative representation for a subset of items. Whether this type of an account is full compatible with extant data depends upon a more detailed exposition of the position. For example, if one accepts the Wickens et al. (1981; 1985) and Halford et al. (1988) assumption that proactive interference has its locus in retrieval alone, there is no obvious explanation in a retrieval cue account for the absence of proactive interference for small set sizes when tested without an interpolated distractor task.

### *Span of Active Processing*

The current study extends the results of prior SAT analyses of item recognition by demonstrating that several items can be associated with a retrieval advantage. This in turn is consistent with the notion that several items can remain active in awareness when there is an intrinsic structure to the material in WM. A natural question concerns whether there is an upper bound on the number of items that can be actively processed in one period of

time. Would, for example, a retrieval advantage be observed for category sizes greater than three?

The present study suggests that what remains active depends on chunking or group processes. While it is certainly possible that more than three items within a category can be simultaneously active in awareness, the upper bound on active processing is likely to depend on the type of structure inherent in the to-be-remembered material and how that structure interacts with particular encoding operations. Additionally, it is reasonable to expect that the upper bound may show substantial variation across individuals (e.g., Conway & Engle, 1994; Ericsson & Pennington, 1993; Halford et al., 1988). Further work is needed to systematically explore these issues.

## APPENDIX

### *Materials*

Animals: dog, horse, pig, cat, cow, bear, wolf, mule, fox, deer, goat, lamb.

Colors: blue, green, white, red, black, pink, brown, gold, beige, gray, tan, mauve.

Body Parts: leg, head, foot, arm, eye, nose, ear, toe, mouth, hand, neck, knee.

Weapons: knife, bomb, sword, gun, club, whip, spear, axe, lance, bow, chain, pipe.

Natural Formations: hill, lake, cave, rock, cliff, plain, ridge, gorge, stream, creek, pond, sea.

Weather Conditions: rain, hail, storm, snow, sleet, wind, fog, ice, smog, clouds, mist, frost.

Clothing: shirt, pants, blouse, socks, shoes, skirt, hat, slip, dress, tie, coat, belt.

Vehicles: car, train, boat, bus, truck, bike, jeep, cab, ship, jet, van, cart.

Insects: fly, bee, wasp, ant, roach, gnat, moth, bug, lice, flea, tick, leech.

## REFERENCES

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Atkinson, R. C., Herrmann, D. J., & Westcourt, K. T. (1974). Search processes in recognition memory. In R. L. Solso (Ed.), *Theories in cognitive psychology*: *The Loyola Symposium* (pp. 101–146). Hillsdale, NJ: Erlbaum.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (1993). Working memory or working attention? In A. Baddeley & L. Weiskrantz (Eds.) *Attention, selection, awareness, and control. A tribute to Donald Broadbent* (pp. 152–170). Oxford: Oxford University Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.) *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.
- Baddeley, A. D., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, **36**, 233–252.
- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, **80**, 1–46.
- Chandler, J. P. (1969). Subroutine STEPIT—finds local minimum of a smooth function of several parameters. *Behavioral Science*, **14**, 81–2.
- Clark, S. E. (1992). Word frequency effects in associative and item recognition. *Memory and Cognition*, **20**, 231–243.
- Conway, A. R. A., & Engle, R. W. (1994). Working—memory capacity as long—term memory activation: An individual—differences approach. *Journal of Experimental Psychology: General*, **123**, 354–373.
- Cowan, N. (1993). Activation, attention, and short-term memory. *Memory and Cognition*, **21**, 162–167.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.
- Doshier, B. A. (1976). The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology*, **8**, 291–310.
- Doshier, B. A. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff or reaction time. *Acta Psychologica*, **43**, 347–359.
- Doshier, B. A. (1981). The effect of delay and interference: A speed-accuracy study. *Cognitive Psychology*, **13**, 551–582.
- Doshier, B. A. (1982). Sentence size, network distance and sentence retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **8**, 173–207.
- Doshier, B. A. (1984). Degree of learning and retrieval speed: Study time and multiple exposures. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **10**, 541–574.
- Doshier, B. A. and McElree, B. (1992). Memory search: Retrieval processes in short-term and long-term recognition. In Larry Squire (Ed.), *Encyclopedia of learning and memory* (pp. 398–406). New York: Macmillian.
- Doshier, B. A., & Rosedale, G. (1989). Combined retrieval

- cues as a mechanism for priming in retrieval from memory. *Journal of Experimental Psychology: General*, **118**, 191–211.
- Doshier, B. A., McElree, B., Hood, R. M., & Rosedale, G. R. (1989). Retrieval dynamics of priming in recognition memory: Bias and discrimination analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 868–886.
- Engle, R. W. (1996). Working memory and retrieval: An inhibition resource approach. In J. Richardson, R. Engle, L. Hasher, R. Logie, E. Stoltzfus, & R. Zacks (Eds.), *Working memory and human cognition* (pp. 89–119). Oxford: Oxford University Press.
- Ericsson, K. A., & Pennington, N. (1993). The structure of memory performance in experts: Implications for memory in everyday life. In G. M. Davis & R. H. Logie (Eds.) *Memory in everyday life* (pp. 241–272). Amsterdam: North Holland.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1–67.
- Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 846–858.
- Gronlund, S. D., Edwards, M. B., and Ohrt, D. D. (1993). *Item information precedes position information in recognition*. Paper presented at the annual meeting of the American Psychology Society, Chicago, Ill.
- Hacker, M. J. (1980). Speed and accuracy of recency judgments for events in short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **6**, 651–675.
- Halford, G. S., Maybery, M., & Bain, J. D. (1988). Set-size effects in primary memory: An age-related capacity limitation? *Memory & Cognition*, **16**, 480–487.
- Hintzman, D. L. (1984). MINERVA2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, **16**, 96–101.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528–551.
- Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, **33**, 1–18.
- Hockley, W. E. (1984). Analysis of response time distribution in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **10**, 598–615.
- Hockley, W. E., & Corballis, M. C. (1982). Tests of serial scanning in item recognition. *Canadian Journal of Psychology*, **36**, 189–212.
- Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model-comparison approach*. San Diego: Harcourt Brace.
- Luce, R. D. (1986a). *Response time: Their role in inferring elementary mental organization*. N.Y.: Oxford Press.
- Luce, R. D. (1986b). Response time distribution in memory search: A caution. In F. Klix & H. Hagendorf (Eds.), *Human memory & cognitive capabilities: Mechanisms & performances, Part A*. New York: North-Holland.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: A time-course analysis. *Journal of Memory and Language*, **32**, 536–571.
- McElree, B. (1996). Accessing short-term memory with semantic and phonological information: A time-course analysis. *Memory & Cognition*, **24**, 173–187.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General*, **118**, 346–373.
- McElree, B., & Doshier, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, **122**, 291–315.
- McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 134–157.
- McElree, B., Jacoby, L. J., & Dolan, P. (1996). *Isolating familiarity and recollective retrieval processes: A time-course analysis*. Presentation at the 37th annual meeting of the Psychonomic Society, Chicago, Ill.
- Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review*, **95**, 183–237.
- Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review*, **63**, 81–97.
- Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, **10**, 465–501.
- Murdock, B. B., Jr. (1971). A parallel-processing model for scanning. *Perception and Psychophysics*, **10**, 289–291.
- 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. (1983). A distributed memory model for serial order information. *Psychological Review*, **90**, 316–338.
- Murdock, B. B., Jr. (1993). TODAM2: A model for the

- storage and retrieval of item, associative, and serial-order information. *Psychological Review*, **100**, 183–203.
- Murdock, B. B., Jr., & Dufty, P. O. (1972). Strength theory and recognition memory. *Journal of Experimental Psychology*, **94**, 284–290.
- Muter, P. A. (1979). Response latencies in discriminations of recency. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **5**, 160–169.
- Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monographs*, **79**(1, Pt.2).
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59–108.
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, **86**, 446–461.
- Ratcliff, R. (1980). A note on modelling accumulation of information when the rate changes over time. *Journal of Mathematical Psychology*, **21**, 178–184.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling the accumulation of partial information. *Psychological Review*, **95**, 238–255.
- Ratcliff, R., & Murdock, B. B., Jr. (1976). Retrieval processes in recognition memory. *Psychological Review*, **83**, 190–214.
- Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, **181**, 574–576.
- Reed, A. V. (1976). The time course of recognition in human memory. *Memory and Cognition*, **4**, 16–30.
- Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors*, **30**, 539–566.
- Seamon, J. G., & Wright, C. E. (1976). Generative processes in character classification: Evidence for a probe encoding set. *Memory & Cognition*, **4**, 96–104.
- Sternberg, S. (1966). High speed scanning in human memory. *Science*, **153**, 652–654.
- Sternberg, S. (1975). Memory-scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, **27**, 1–32.
- Tehan, G., & Humphreys, M. S. (1995). Transient phonemic codes and immunity to proactive interference. *Memory & Cognition*, **23**, 181–191.
- Tehan, G., & Humphreys, M. S. (1996). Cuing effects in short-term recall. *Memory & Cognition*, **24**, 719–723.
- Theios, J. (1973). Reaction time measurement in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7, pp. 44–85). New York: Academic Press.
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. New York: Cambridge University Press.
- Wickelgren, W. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, **41**, 67–85.
- Wickelgren, W. A., & Norman, D. A. (1966). Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, **3**, 316–347.
- Wickelgren, W. A., Corbett, A. T., & Doshier, B. A. (1980). Priming and retrieval from short-term memory: A speed-accuracy tradeoff analysis. *Journal of Verbal Learning and Verbal Behavior*, **19**, 387–404.
- Wickens, D. D., Moody, M. J., & Dow, R. (1981). The nature and timing of the retrieval process and of interference effects. *Journal of Experimental Psychology: General*, **110**, 1–20.
- Wickens, D. D., Moody, M. J., & Vidulich, M. (1985). Retrieval time as a function of memory set size, types of probes, and interference in recognition memory. *Journal of Experimental Psychology*, **11**, 154–164.

(Received March 12, 1997)