Reprinted from:
SHORT-TERM MEMORY

ACADEMIC PRESS, INC.
New York San Francisco London

CHAPTER

9

DYNAMICS OF RETRIEVAL[1]

WAYNE A. WICKELGREN

When a subject attempts to recall or recognize many items or pairs of items over an extended period of time (15 sec or more), the retrieval

[1] This work was supported by Grant MH 17958 from NIMH.

process appears to be extremely complicated and it is difficult to describe the dynamics of such complex retrieval processes with mathematical precision. For this reason, complex retrieval tasks, such as free recall and ordered recall of entire lists, will be largely ignored in this paper.

This paper will try to draw some tentative conclusions concerning the retrieval processes in what might be called "elementary" recall and recognition tasks. Tasks that probe for recall or recognition of a single item or pair of items in a short period of time (less than 5 sec) are assumed to be tapping elementary retrieval processes that can be characterized in a precise and simple theory. According to this hypothesis, complex retrieval tasks, such as those involved in free recall or everyday recall of large amounts of information, are composed of sequences of elementary retrieval processes plus logical or other cognitive operations being applied to draw deduction from these retrieved memories.

Three elementary retrieval processes will be assumed: recognition, recall, and recency judgments. This paper defends the direct-access strength theory of elementary recognition and recall. According to this theory, presentation of a single item in a recognition test elicits a feeling of familiarity for the item as associated to a particular experimental context; the greater the strength of association between the item and the context, the greater the feeling of familiarity that is elicited. If a pair of items is presented, the greater the strength of association between the pair of items, the greater the feeling of familiarity that is elicited.

Following the criterion decision rule of Thurstonian Scaling and signal-detection theory, the subject is assumed to say "yes" to a test item or pair in a recognition task, if and only if the strength of that item or pair exceeds a variable recognition criterion. In multiple-choice recognition memory or recall with a small number of alternatives, an individual is assumed to choose the alternative that has the greatest associative strength in memory. In recall, with an extremely large number of alternative responses, let us assume that the correct response will be selected only if its associative strength exceeds some (high) recall threshold.

Elsewhere, I have speculated that some recency judgments are based on a second property of the long-term memory trace, namely, its resistance, which increases monotonically with delay since learning (Wickelgren, 1972). However, recency judgments can clearly be mediated indirectly by recalling associations to time concepts (serial position, times of the day, dates, periods in one's life, *etc.*). Because comparatively little research has been done on recency memory, and that research has not generally attempted to rule out indirect mediation of recency judgments via associations to time concepts, little can be concluded regarding the reality of any elementary recency retrieval process, let alone determining its logical character and dynamics. For this reason, this chapter will be concerned only with analyzing the dynamics of elementary recall and recognition.

Recently there has been some interest in comparing recognition with free recall. Such comparisons are of little theoretical interest because free recall is certain to be a very complex task involving a complicated sequence of elementary recall and recognition retrieval operations over a considerable period of time. Also, experimental control of the cues used in free recall is typically very poor, and subjects must generate most of the cues (typically category labels) themselves if they are to recall any reasonable number of items. There are bound to be many differences between the retrieval processes involved in recognition and free recall, and such comparisons tell us very little concerning the precise nature of the retrieval processes involved in elementary recognition or recall. The more illuminating comparisons are of probe recognition and probe recall with retrieval times limited to a few seconds.

I. STRENGTH VERSUS SCANNING THEORIES OF RETRIEVAL

At least three major stages of the memory usage process need to be distinguished: (1) the time to perceive the test stimulus, (2) the time to retrieve memory traces in some way connected to the trace stimulus, and (3) the time to make a response. The memory-retrieval phase of this process is sometimes divided into two subphases: (a) the time to access one or more memory traces and (b) the time to make a decision, or choose a response, based on the memory traces that have been accessed.

There are two basic classes of retrieval theories: strength theory (Baddeley & Ecob, 1970; Corballis, Kirby, & Miller, 1972; Norman & Wickelgren, 1969; Okada, 1971; Wickelgren & Norman, 1966) and scanning theory (Sternberg, 1966, 1967, 1969). The theory of Juola, Fischler, Wood, & Atkinson (1971) incorporates a mixture of strength and scanning theories.

Strength theory assumes that a test item directly accesses its internal representation in memory without the need for memory scanning, or search. Associated with this internal representative is a memory-trace strength that provides the input to the decision-making system, which selects the response. In addition, strength theory postulates that the output of the memory-accessing process is a continuous real-valued variable (strength), rather than a discrete (all-or-none, finite-state) variable. A continuous speed–accuracy tradeoff in memory retrieval can be realized within a strength theory, if the output of the memory-accessing processes increases in some continuous way with the time allowed for retrieval.

Memory-scanning theories assume that memory accessing involves a search of a number of storage locations, looking for a match between a stored item and the test item. In recognition, the output of the memory-accessing process is a one–zero (all-or-none) variable indicating the presence or absence of a match between the test item and an item in some storage location. In probe recall, after a match is found, the subject recalls the item in the next storage location. Since, in general, several memory locations are searched, a vector of ones and zeros will be output from the memory-accessing process as input to the decision process. The memory-search process is almost always assumed to be sequential, and the output of the process is assumed to be a series of discrete (zero *vs.* one) variables. Although one could postulate the output of the matching process to be a vector of continuous variables, representing the degrees of similarity between the test items and the items in storage locations, this sort of hybrid model has never, to my knowledge, been proposed. A scanning theory could account for continuous speed–accuracy tradeoff in retrieval by assuming that the scanning rate was variable and that, at faster scan rates, the probability of an error occurring in matching was higher than at slower scan rates.

Although the memory-scanning theory achieved almost instant popularity following the initial Sternberg (1966) study, and many subsequent studies concerned with the dynamics of memory retrieval have presented interpretations of their results in terms of scanning theory, the accumulated evidence strongly favors some form of strength theory over any form of scanning theory for both recall and recognition. The remainder of this section is concerned with documenting this assertion.

A. Recognition Memory

1. Short-Term Memory

Sternberg (1966) investigated reaction time in probe recognition memory with target sets in short-term memory consisting of from 1 to 6 digits. The task for a subject was to decide whether or not the test digit was a member of a previously memorized list (target set) consisting of from some 1 to 6 digits. Sternberg found that the time to make this decision increased linearly with the length of the target list, with the same slope constant obtaining for both negative and positive responses. Furthermore, Sternberg found no significant differences in reaction time as a function of serial position within a list of a given length. From these results, Sternberg concluded that the access process in recognition memory retrieval was an exhaustive serial scan of the target list. That is to say, subjects were assumed to scan every item in the target list for a match with the test item and make their yes–no decision on the basis of whether any match was achieved after scanning the entire list. It might seem more natural to assume that the scan would terminate when a match was achieved, but the scanning theory appears to require the assumption that all items in the target list are scanned, primarily in order to account for the fact that the slope for negative and positive items was identical in the Sternberg study. Although none of these results is inconsistent with strength theory, any strength theory accounting for these results would be relatively *ad hoc* and would not account for the results in the elegant manner provided by the exhaustive-scanning theory.

Although Sternberg's results have been replicated many times when there has been strict adherence to certain critical aspects of experimental procedure, it is also quite clear that Sternberg's findings are not generally true under other experimental procedures. The results under these other conditions favor strength theory over scanning theory. Furthermore, it is clearly the case that insufficient attention has been paid to the problem of speed–accuracy tradeoff in retrieval in those experiments alleged to support the scanning theory. When attention is paid to this problem, the results in those studies (*e.g.*, Sternberg, 1966) may not support memory-scanning theory either. Finally, there is evidence using the standard Sternberg (1966) conditions that is contrary to memory-scanning theory (Baddeley & Ecob, 1970).

Basically, the Sternberg results are replicated when presentation rate for the elements in the target set is relatively slow (about 1 sec per item or slower), and a relatively long (2 sec) interval elapses between the end of the list and the presentation of the test stimulus. Slow presentation rate and substantial delays between list presentation and test allow ample opportunity for rehearsal. Such rehearsal could easily eliminate the strong recency effect and any differences in degrees of learning (primacy effects) that would be expected by a strength theory. Consistent with this interpretation, when presentation rates are faster and delays between presentation and test are shorter, strong serial-position effects are obtained that are dominated by the recency effect, precisely as expected by strength theory (Baddeley & Ecob, 1970; Burrows & Okada, 1971; Corballis, 1967; Corballis *et al.*, 1972; Kirsner & Craik, 1971; Morin, DeRosa, & Stulz, 1967; Morin, DeRosa, & Ulm, 1967). In addition, Okada (1971) found a monotonic increase in "hit" latency with increasing lag in a continuous-recognition memory task.

Although only a few studies actually reject the original Sternberg (1966) finding of a linear increase in recognition-memory reaction time with increasing list length, many other studies show gross but not significant deviations from linearity. Furthermore, in every study known to me in which error rates have been reported, the error rates increase monotonically (and usually substantially from a relative point of view) with list length. If there is some type of speed–accuracy tradeoff operating in recognition memory, then such results indicate that the recognition-memory latencies at long list lengths are underestimated. In the absence of any knowledge concerning the form of speed–accuracy tradeoff, we cannot know whether this is a major or minor effect. However, at some point, we must assume that only small increases in accuracy occur with increasing decision latency. That is to say, after some time, it may require a relatively large amount of additional decision time in order to achieve a relatively small increase in recognition-memory accuracy. If only some of the decision times found in the memory-scanning studies lie in this range, small differences in recognition-memory accuracy could translate into very large differences in recognition time. This would mean that the obtained reaction times for long list lengths are *grossly* underestimated in relation to the reaction times for short list lengths.

The mistake is frequently made in interpreting error differences as a function of list length that a difference in error rate of 1% *vs.* 3%, or 0.1% *vs.* 1%, is a small difference in recognition memory accuracy. This is surely false from the point of view of strength theory, though it could be true with some memory-scanning theory. However, the general point is we cannot know at present what differences in recognition memory accuracy are "small" and insignificant in their biasing effects on retrieval time.

Also, the dynamic range of the independent variable (target set size) is so small in most studies (generally less than a range from 1 to 6 or 2 to 6 items), that a large variety of different functions will provide a reasonable fit to the increase in recognition memory-reaction time as a function of list length.

For all these reasons, there is little reason to have confidence in the alleged linear functions relating short-term recognition-memory time to list length.

Finally, even in the standard Sternberg conditions for a short-term memory recognition time experiment with slow presentation rate (1.2 sec per digit) and a 2 sec delay interpolated between presentation of the list and presentation of the test item, Baddeley and Ecob (1970) found that increasing the frequency of presentation of digits in the target list reduced the recognition-memory reaction-time for these digits by comparison to nonrepeated digits in the same target list and by comparison to comparable items in lists without repeats. These results are exactly what would be expected by a strength theory and are clearly inconsistent with the scanning theory. The inconsistency with a scanning theory is especially dramatic, since in one of the Baddeley and Ecob experiments, a linear increase in reaction time with increasing list length was obtained accompanied by an absence of serial-position effects. This demonstrates that the principal results claimed as support for the scanning theory can be obtained under conditions where a strength theory clearly appears to be necessary in order to account for the results.

2. Long-Term Memory

As Baddeley and Ecob (1970) and Corballis *et al.* (1972) point out, it was never reasonable to imagine that a serial-scanning process operated as a general model of the recognition-memory process. To apply the memory-scanning theory to long-term recognition memory, one must, at a minimum, assume that subjects can directly access some relevant list of location to scan, as it is obviously false to assume that we scan all our long-term memory-storage locations at even the "high-speed" scanning rates found by Sternberg (1966). Long-term recognition memory would simply take far longer than it does take.

The most favorable context for evaluating the scanning theory in retrieval from long-term memory is that of deciding whether a target word is a member of a verbal category (bird, animal, woman's name, *etc.*). In these studies, it has been shown that recognition-memory time generally increases with the number of instances in the category (Juola & Atkinson, 1971; Landauer & Freedman, 1968; Meyer, 1970; Wilkins, 1971). But there have also been two failures to find any significant effect of category size (Collins & Quillian, 1970, Experiment I; Egeth, Marcus, & Bevan, 1972). The two failures come from radically different experimental contexts and represent gross extremes in the manipulation of category size. Thus, we cannot conclude that category size has any effect on retrieval time from long-term memory. However, even if this conclusion could be substantiated, it is undoubtedly predicted by both scanning and strength theories, since the strength of an association between a category and an instance of a category very likely decreases on the average for categories with larger numbers of examples.

Furthermore, there are discrepancies from any simple scanning model apparent in the data for both recognition-memory retrieval from well-established long-term categories (Juola & Atkinson, 1971) and categories newly established by presentation of a group of instances in the same experimental context (Fischler & Juola, 1971; Juola *et al.*, 1971). The principal discrepancies in the case of the experimentally established "categories" are recency effects (items presented more recently have shorter recognition-memory reaction times) and frequency effects (positive items presented more frequently have shorter recognition times than items presented less frequently and distractors presented more frequently have longer recognition latencies than distractors presented less frequently). These effects are, of course, exactly what a strength theory would expect. Faster recognition times for more recent and more frequent items have also been found in a continuous recognition memory task by Hintzman (1969).

Thus, despite an enormous body of work done from the perspective of memory-scanning theory, the results, at present, unequivocally support a strength theory over a scanning theory for recognition-memory retrieval time for both long-term and short-term memory.

B. Recall

1. *Short-Term Memory*

The dynamics of retrieval via a recall test have not been as extensively or systematically investigated as for recognition tests. Nevertheless, it can be said that the results on the whole are somewhat more consistent with the strength theory than with the scanning theory of retrieval in recall.

The scanning theory proposed by Sternberg (1967) for recall postulated a self-terminating serial scanning, rather than an exhaustive serial scanning (as in the case of recognition memory). The basic reason Sternberg postulated self-terminating scanning was that recall time increased monotonically (and approximately linearly) with position in the list, the fastest times being for the first position and the slowest times being for the last position in the list. However, it was clear from the data that one could not assume that all subjects were beginning their scan at the beginning of the list, since the time to recall the second item, given the first item as a probe, increased as a function of list length. Furthermore, the slope of the function plotting recall time against serial position was not equal to twice the slope of the function plotting recall time against list length, as it should have been according to a self-terminating serial-scanning strategy.

To resolve these discrepancies, Sternberg proposed that subjects sometimes begin at other positions in the list and scan cyclically. In support of this interpretation, the results for two different subjects were plotted with the different subjects showing vastly different rates of increase in recall time as a function of serial position. Sternberg claims that the average recall data reflect a mixture of starting strategies from different subjects.

However, it is at least equally plausible to assume that under the conditions used by Sternberg (which encouraged extensive rehearsal of the list between presentation and test), strengths of associations decreased systematically as a function of serial position. Assuming that recall time decreases with increasing associative strength, the Sternberg results can be accounted for by a strength theory.

Once again, the reality of the linear increase in recall time as a function of list length is very questionable considering the fact that errors in recall increase markedly as a function of list length. As mentioned previously, consideration of the possibility of speed–accuracy tradeoff in recall would indicate that under these conditions, the recall times for long lists are underestimated relative to the recall times for short lists. Whether consideration of the matter would make the form of the function nonlinear is unknown, but it is a possibility.

The monotonic increase in recall latencies as a function of serial position (pure primacy effect) is very rarely observed in either short-term or long-term recall experiments. A straight recency effect was found in short-term recall times by Norman (1966), and both primacy and recency effects on recall time have been found in a task that presumably involves a combination of short- and long-term memory by Waugh (1970) and in a long-term memory task by Kennedy (1968). DeRosa and Baumgarte (1971) find both primacy and recency effects on recall time and also find effects on recall time induced by inserting, in the middle of the list, brief pauses that induce a grouping structure. These enormous variations of recall time as a function of serial position are essentially equally easy or equally difficult for either memory scanning or strength theory to account for, but they indicate how little can be concluded from the serial-position results obtained by Sternberg (1967).

2. *Long-Term Memory*

Once again, as was the case for recognition time, greater frequency of presentation of a pair of items decreases the probe-recall latency for the response item given the stimulus item, even for a very large number of trials beyond the last error (Eimas, 1964; Eimas & Zeaman, 1963; Hall, 1969; Millward, 1964; Peterson, 1965; Schlag-Rey, Groen, & Suppes, 1965; Suppes, Groen, & Schlag-Rey, 1966; Theios, 1965; and Wiggins, 1957).

Although frequency and recency effects are the stock-in-trade of direct-access theories (whether finite-state or strength theories), note that very likely some type of scanning theory can be devised to account for such effects, but it would have to be a great deal more complex than the scanning theories proposed so far.

The variable of primary significance according to a scanning theory is the size of the category through which the scan takes place. In the case of long-term recall, it is possible to find conditions, such as in Metlay, Handley, & Kaplan (1971), in which subjects do appear to search sequentially through the instances of a category name, such as the 12 zodiac names, looking for a name containing the letter "n" (*e.g.*, Gemini, Cancer, and Capricorn). However, Freedman and Loftus (1971) have shown that the number of examples in a category has no necessary effect on recall reaction times when the cues are more strongly associated to correct response items, for instance, producing an animal name that starts with the letter "z," such as "zebra." For tasks such as these, where both the category name and the initial letter are strongly associated to a target word or words, there appears to be no effect of either the number of instances in the noun category or the number of possible correct instances. Such findings seem strongly inconsistent with serial scanning as a general model of retrieval in recall from long-term memory.

That the strength of association between the cues and the response items constitutes the principal factor in determining whether a memory-search process needs to be initiated was also indicated by the results of Ceraso, Bader, and Silverstein (1970), who found that the effects of list length on response latency in paired-associate recall declined with increasing degree of learning of the lists.

On the whole, it would seem that the evidence favors some type of strength theory of retrieval in recall as well as in recognition memory.

II. SPEED–ACCURACY TRADEOFF IN MEMORY RETRIEVAL

As was mentioned in the preceding sections, our lack of knowledge concerning the possibility of speed–accuracy tradeoff in memory retrieval limits our ability to draw conclusions about retrieval dynamics on the basis of recognition and recall times. Stating the issue in strength-theory terms, the issue is whether the strength of the accessed (retrieved) memory trace increases in a gradual way as a function of retrieval time or suddenly appears in an all-or-none manner at some fixed time following the initiation of the retrieval process. As usual, it does not appear to be possible at present to distinguish a gradual incremental increase in retrieved strength as a function of retrieval time from an all-or-none retrieval process with some appropriate probability density function for finishing times. Each would produce a continuous speed–accuracy tradeoff function, and these two theories will be lumped together and considered in opposition to a theory in which retrieval time is relatively fixed.

If retrieved-memory strength varies in a relatively continuous manner with retrieval time, then presumably one could obtain a family of functions such as that shown in Fig. 9-1.

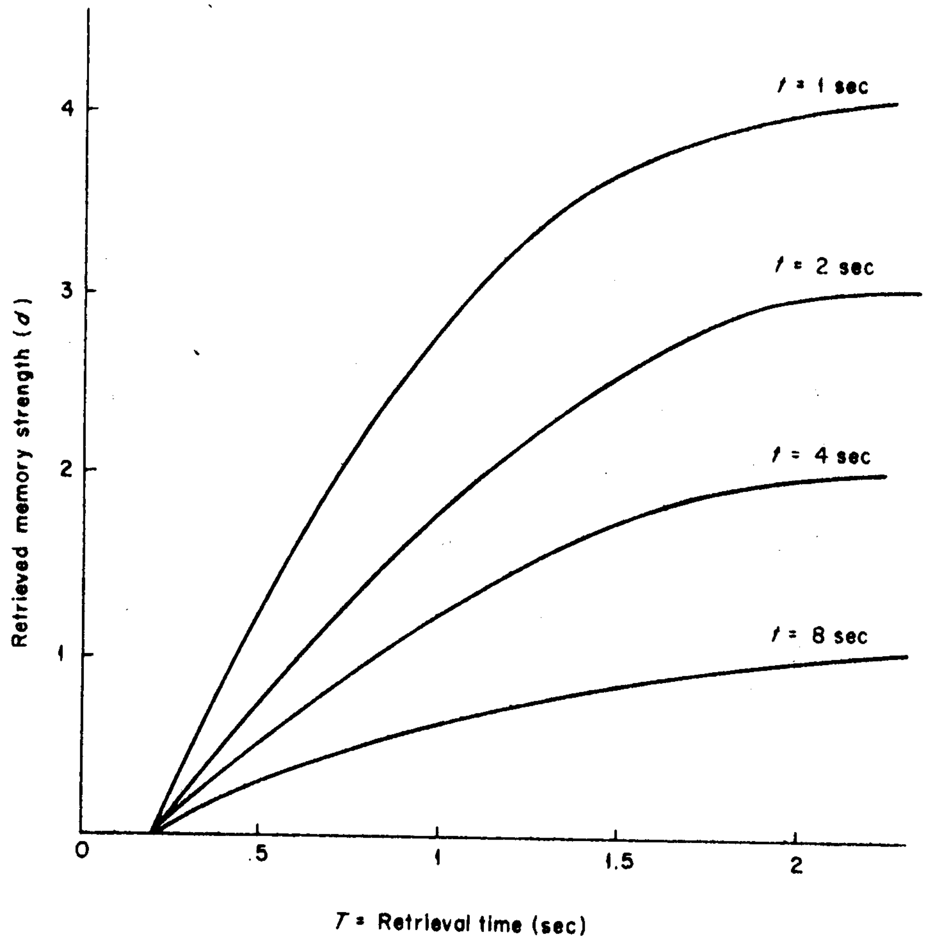

Fig. 9-1. Hypothetical speed–accuracy tradeoff functions for memory retrieval in a short-term memory task with retention interval (t) as the parameter.

The hypothetical speed–accuracy tradeoff function shown in Fig. 9-1 indicates the possibility of an initial latent period of perhaps 200 msec in which the memory-retrieval process would not produce any strength output. After this, the retrieved-memory strength is presumed to increase as a function of retrieval time (T), at a monotonically decreasing rate, to an asymptote set by the strength of the memory trace in store at that time. In Fig. 9-1 this is shown as a function of retention interval (t) in a short-term memory task, where the strength of the trace in storage is presumed to decrease markedly over a retention interval from 1 to 10 sec. Previous studies deriving the strength of the memory trace from recognition or recall-accuracy (choice) data presumably yield the asymptotic values of the function shown in Fig. 9-1. That is to say, if subjects are allowed at least 3 or 4 sec for retrieval, they are presumed to be at or near the asymptote in terms of retrieved strength as a function of retrieval time. Some evidence in support of this assumption will be discussed later.

A. Methods Used to Determine Speed–Accuracy Tradeoff Functions

In determining a speed–accuracy tradeoff function for recognition memory retrieval such as that shown in Fig. 9-1, it is important that the reaction time be an independent (manipulated) variable rather than a dependent (measured) variable. Subjects should be induced to respond at different speeds in different sessions or different blocks of the same experimental session. This procedure for inducing subjects to respond at different speeds might involve the use of different payoff matrices for correct vs. incorrect responses emitted at different response times, although this procedure often appears to encourage subjects to make fast *random* guesses.

Alternatively, one could use simple instructions to the subjects to respond in different time windows or before some deadline time with the window or the deadline time being varied across different sessions or blocks of a session. In using either procedure, it is probably helpful to provide subjects with some feedback concerning their reaction time following each trial.

In my opinion, the best method of manipulating response time is to provide an external cue telling the subject when to respond. This is the method used by Schouten and Bekker (1967), and it appears to produce response-time distributions with very low variance. Low variance response-time distributions minimize the distortions of the speed–accuracy tradeoff function that are produced by pooling results over large sections of nonlinear functions (such as the functions in Fig. 9-1).

Fortunately, perceptual speed–accuracy tradeoff functions have primarily been determined using manipulated response times (see Pachella & Fisher, 1972, Fig. 2; and Schouten & Bekker, 1967, Figs. 2, 3, 4, and 6).

However, recently, several papers have appeared in the perceptual speed–accuracy area (Lappin & Disch, 1972a, and 1972b; Rabbitt & Vyas, 1970) in which speed–accuracy tradeoff functions (called latency operating characteristics by Lappin & Disch) have been derived using measured reaction times instead of manipulated reaction times. The procedure involved determining the accuracy measure for responses emitted with reaction times lying within some interval and doing this for all intervals of reaction time for which a sufficient number of responses were obtained in the experiment. Then accuracy is plotted as a function of the mean or median reaction time for each interval. This procedure is theoretically deficient for the following reason: The responses selected within each reaction-time interval do not necessarily arise from the same underlying sensory (or memory) strength distribution. It is reasonable to suppose and indeed many models of perceptual or memory reaction time assume that events with greater strength will be responded to more rapidly than events with lesser strength. To the extent that strength under a constant experimental condition is a random variable with some substantial variance, a "constant" experimental condition will generate a distribution of strength values. According to the present conception, the trials on which the condition has greater strength will initiate responses at shorter latencies on the average than trials with lesser strengths.

Direct evidence for the validity of this hypothesis has been provided for the recognition memory area by Norman and Wickelgren (1969), who showed that yes-no responses that were subsequently given high confidence were initiated with short latencies. In the presence of this confounding factor, the effect of plotting accuracy against measured reaction time is to overestimate the accuracy at short reaction times and underestimate the accuracy at long reaction times by comparison to a procedure that generates speed–accuracy tradeoff functions by manipulating rather than measuring reaction times.

Schouten and Bekker (1967) obtained speed–accuracy tradeoff functions using both procedures. What they called the "free" procedure used measured reaction times, and the "forced" procedure used manipulated reaction times. Comparing their Figs. 1 and 2 demonstrates a huge overestimation of accuracy at short latencies using measured reaction times, confirming the present argument. There was no possibility of seeing the predicted underestimation at longer latencies in their task because performance became perfect at long latencies (floor effect).

B. Universal Retrieval Function?

It would be particularly simple if it could be shown that the speed–accuracy tradeoff functions for memory retrieval could always be factored into two functions: one expressing the level of the asymptote as a function of various experimental conditions (such as learning time and retention interval) and the other function expressing the approach to this asymptote as a function of retrieval time. For the family of functions shown in Fig. 9-1. such factoring could be given by the following equation: $d(t, T) = m(t) \cdot r(T)$, where $m(t)$ represents the strength of the memory trace in storage as a function of retention interval and $r(T)$ represents the universal memory-retrieval function. Of course, there might be different memory-retrieval functions for recall and recognition, different modalities, and as a function of who knows what other conditions. This discussion is meant to indicate only the possibility that speed–accuracy tradeoff functions could have a very simple lawful character.

One particularly simple and plausible form for the "universal" retrieval function might be: $r(T) = 1 - e^{-\gamma[T-\delta]}$, meaning an exponential approach to a limit following a lag of δ, with $[T - \delta] = T - \delta$ for $T > \delta$ and $[T - \delta] = 0$ for $T \leq \delta$. Different retrieval functions for recall *vs.* recognition, short- *vs.* long-term memory might be simply characterized by different values of the parameters, γ and δ.

C. Incremental Tradeoff versus Fast-Guess Theory

Of course, one could obtain empirical curves similar to those shown in Fig. 9-1 without there being a continuous increase in retrieved memory strength as a function of retrieval time, but rather as a result of a changing mixture of fixed (long latency) accurate responses and fast guesses with chance accuracy. In perceptual reaction time contexts, some results have favored a pure fast-guess theory (Swensson & Edwards, 1971; Yellott, 1971), and some results have favored a continuous speed–accuracy tradeoff (Swensson, 1972).

There is little doubt that subjects can make fast chance guesses. The really important question is whether subjects can adjust to retrieval times in some manner that yields incremental growth of accuracy. If it is possible to trade off accuracy for speed in a relatively continuous manner over some range of retrieval times, then this is a critical aspect of memory dynamics for which the laws have yet to be determined. If this is not possible, it would be important to know that it is not possible.

Several previous studies concerned with speed–accuracy tradeoff in memory retrieval suffer from the possibility of interpretation purely in terms of the fast-guess theory and do not unequivocably demonstrate continuous speed–accuracy tradeoff in either recognition memory (*e.g.*, Lively, 1972; Swanson & Briggs, 1969) or recall (Murdock, 1968).

However, a recent study by Reed (1973) obtained continuous speed–accuracy tradeoff in recognition memory, using a method somewhat similar to that used by Schouten and Bekker (1967), in which the results precluded interpretation in terms of fast guesses. The results were consistent with an exponential growth to a limit, though Reed favors a different theoretical function.

Reed's results indicate that the recognition retrieval function for short-term memory is close to its asymptote by about 2 sec after the onset of the test item. This agrees with the results of Murdock (1968) for short-term recall. Thus, we can probably assume that all the choice data obtained with retrieval times of 3 or 4 sec or more are obtained from the asymptotic section of the retrieval function and thus are relatively insensitive to latency differences. Both Reed and Murdock found the bulk of the increase in memory accuracy within the first second, indicating that this is the primary region of interest in investigating speed–accuracy tradeoff in (short-term) memory. Finally, the weak qualitative agreement of the Reed and Murdock studies suggests, very tentatively, the possibility that the retrieval dynamics of short-term recall and recognition memory are identical.

III. RETRIEVAL INTERFERENCE

Response Competition and Blocking

For many decades now it has been recognized that forgetting as measured by the increasing inability to produce the B member of an AB paired associate is not due only to decreasing strength of the association from A to B. Learning a list of AB paired associates followed by learning a list of AC paired associates creates problems in the subsequent recall of either the first or second list associates of the A stimulus due to factors other than the strengths of the associations from A to B and A to C. If subjects are asked to recall only the first (or the second) associates of the A items, they will show poorer performance following the learning of an AC list than a CD list in part because subjects do not discriminate perfectly which (B or C) response was paired with the A stimulus in the first (or second) list. When only one item may be recalled, retroactive and proactive interference designs that pair two responses with the same stimulus would create competition between the two responses, even if there were no storage interference effects on associative strengths due to learning two responses to the same stimulus.

In modified modified free recall (MMFR) designs, subjects are allowed to produce two responses to a stimulus item. While MMFR is designed to eliminate the logically necessary competition involved in having two responses to the same stimulus, it may not eliminate all interfering effects of an AC association on recall of an AB association. A strong AC associate appears to block recall of an AB associate as evidenced by the fact that actually supplying the second list (C) associate of an A stimulus item depresses recall of the first list (B) associate by comparison to an ordinary MMFR procedure (Postman, Stark, & Fraser, 1968). Also, in free recall of the 50 states in the U.S., prior study of a subset of 25 state names depresses recall of the remaining 25 state names (Brown, 1968;

Karchmer & Winograd, 1971). Thus, recall tests may involve an unconscious blocking of weaker associates by stronger associates, in addition to any conscious response competition.

It seems very reasonable to assume that a recognition test in which one presented both the A and B members of the pair sometimes correctly paired (A_iB_i) and sometimes incorrectly paired (A_iB_j) would get around competition and blocking. In addition to this being reasonable, there is also direct evidence that the presence of a strong AC associate does not diminish the discriminability of correct (A_iB_i) from incorrect (A_iB_j) paired associates in a short-term recognition memory task (Wickelgren, 1967). This independence from irrelevant associations in a recognition task could be demonstrated for short-term memory because learning an AC associate in short-term memory does not appear to depress the strength of an AB associate in storage any more than does learning a CD associate depress the strength of the AB associate. Storage interference in short-term memory appears to be independent of this type of fine-grain similarity (although storage interference in short-term memory is not independent of grosser modality similarity of interpolated learning).

Unfortunately, storage interference in long-term memory does appear to be similarity dependent. See Wickelgren (1972) for a long list of studies supporting the similarity dependence of storage interference in long-term memory. Thus, it is somewhat more difficult to demonstrate the independence from this type of retrieval interference in recognition based on long-term traces. It is possible to test for independence from irrelevant associations in recognition based on long-term memory traces with a PI design, where the AC or CD learning is done prior to the AB learning. AB learning should be to the same criterion in both cases. If long-term recognition tests are also free from associative competition and blocking, then subsequent tests of AB recognition should be identical in both cases. That is, recognition tests should show no PI. Evidence for the absence of PI in recognition matching tests has been obtained by Postman *et al.* (1968), confirming this prediction.

B. Response Suppression

Recently, there has been some evidence to support the notion that one of the factors producing declines on recall tests as a function of time is the loss by a subject of his "set" for the responses. This loss of availability of the responses (possibly active reponse suppression) combines with blocking and losses in specific associative strength to produce the forgetting shown in recall tests (Birnbaum, 1972; Cofer, Faile, & Horton, 1971; Postman *et al.,* 1968).

To my knowledge no one has ever suggested that there could be any analogue of response suppression operative in a recognition-memory task, and it is difficult to see any logical possibility for this. Certainly the availability of the yes–no responses could not in any sense be thought to be reduced over the retention interval. Thus, it is difficult if not impossible to imagine any analogue of this loss of response availability operating as an interference factor in a recognition test.

C. Context Effects

It has sometimes been claimed that cues from the background context become strongly associated to the material being learned, and the presence or absence of these contextual cues at the time of retrieval affects the probability of correct recall or recognition. In discussing the effects of changes in context from the time of learning to the time of retrieval, it is important to make the distinction between changes in "irrelevant" cues and changes in "relevant" cues.

Irrelevant cues refer to cues that are held constant throughout the learning of a list. Irrelevant context factors include such properties as the color of the stimulus items, the color of the ground in which the stimulus items are presented, the sex of the experimenter, various characteristics of the testing room, the time of day, the posture of the subject, the physical nature of the apparatus used to present the stimuli and record responses, the presence or absence of background music or other auditory background, *etc.*

However, if the color of each stimulus item is changed systematically from one paired associate to another with the A_1 item always appearing in red, the A_2 item always appearing in blue, and the A_3 item always appearing in green, *etc.*, then the color of the stimulus item can no longer be considered an irrelevant context cue. This latter case must be considered to be an example of paired-associate learning with compound stimuli, where either the color or the verbal character of the stimulus or both could be used as a specific relevant cue to be associated to the response.

McGeoch and Irion (1952, pp. 448–451) listed a number of studies supporting the importance of contextual change as a retrieval interference factor in memory. However, closer examination of this experimental work and other studies on the effects of changing background contexts reveals only a few studies that find significant effects of altering truly irrelevant background contextual cues. Many studies that have been alleged to show the importance of background context were actually experiments in which compound (relevant) stimuli were changed from learning to test. There is every reason to expect that changing or eliminating one of the relevant attributes of the stimulus would depress memory performance. Most of the studies that involve varying truly irrelevant background context have failed to find significant effects of altering such context. Unpublished work of my own has also failed to reveal significant effects of contextual change on either recall or recognition of Russian–English word pairs, though the effect of contextual change on recall was almost significant. Since some of the alterations in experimental context have been deliberately designed to be quite extreme, the negligible or unreliable effect of the contextual changes argues strongly that irrelevant context often plays little role in the coding and retrieval processes of memory.

However, there has recently been a resurgence of interest in the effects of alterations in background context as a possible retrieval interference factor in recognition memory (DaPolito, Barker, & Wiant, 1971; Light & Carter–Sobell, 1970, Thomson, 1972; Tulving & Thomson, 1971; Winograd & Conn, 1971; Winograd, Karchmer, & Russell, 1971). Although the specific designs of these various studies have differed somewhat, the nature of the context effects found for recognition memory can be illustrated by considering the initial study by Light and Carter–Sobell. In this study, recognition memory for nouns that have at least two very different meanings was studied, for example, the noun "jam" as in "raspberry jam" and as in "traffic jam." Although subjects at the time of the recognition test were only asked to identify whether the noun "jam" had occurred earlier, this noun was accompanied at the time of test by an adjective that biased one or another meaning for jam. This adjective could be the same as the adjective presented with the noun during learning of the list (*e.g.*, "raspberry jam" in both cases), or a different adjective, but one that biased the same meaning of the noun ("strawberry jam" in learning followed by "raspberry jam" in the test), or a different adjective that biased a different meaning for the noun ("traffic jam" followed by "raspberry jam"). Changes in the adjective context lowered the discriminability of presented from unpresented nouns by a substantial amount, though recognition in this and all studies of nouns presented in new semantic contexts is far above chance.

However, once again, there has been a failure to recognize the distinction between changes in irrelevant and relevant cues. In the Light and Carter–Sobell study, subjects were not informed at the time of learning that the adjectives were irrelevant and, in any event, the adjectives were different from one noun to another, which would certainly have caused them to have been learned by the subjects and strongly associated to their respective nouns. Thus, this is a study of the effects of varying a relevant part of a compound stimulus, which is logically distinguishable from varying irrelevant context. The same criticism applies to all the other studies and thus these studies provide no support for the role of irrelevant context in recognition memory. All such studies show is what we already knew: that very frequently two cues are better than one. That is, if one learned two relevant cues, then the presence of both cues at the time of the retention test will enhance performance as compared to a condition where only one relevant cue is available.

However, this work does support the hypothesis that, to some extent, what we associate in our minds are the internal representatives of the *meanings* of words (concepts), in addition to, or instead of, associating the internal representatives of *words,* though there does also appear to be a substantial word-strength component to memory in these tasks (Davis, Lockhart, & Thomson, 1972). Also, it is now well established that mnemonics based on the meanings of words in a pair can substantially improve one's learning and retention of paired associates. Thus, this work raises the question concerning what would happen if a subject thought of a different meaning for a word at the time of the test from the meaning thought of at the time of learning. To the extent that some of the memory trace is tied to the specific meaning and not available at the time of retrieval if a different meaning is thought of, there should be a depression of recognition-memory performance. The studies cited earlier demonstrate that this can occur under special circumstances, though I hope it is clear from the preceding discussion that these effects of altering compound stimuli do not prove anything concerning the effects of small changes in irrelevant context in ordinary list-learning tasks. It may well be that the probability that a subject thinks of the same meaning or sets of meanings at the time of retention as he did at the time of original learning is so high in most learning and memory tasks that these potentially significant retrieval interference effects can, in practice, be completely ignored.

At present, we do not know what portion of the forgetting shown by recognition tests is due to a change in the representation of the stimulus item as a function of retention interval and what portion is due to a change in the strength of association from the representation of the stimulus to the representation of the response as a function of retention interval. Also, it seems perfectly reasonable, at present, to imagine that a change in the representation in the stimulus item is as much a storage interference factor as is a change in the association from the representation of the stimulus to the representation of the response.

If the probability of interpreting the stimulus in a different manner at retrieval from the interpretation given at the time of learning is constant as a function of retention interval (as it is in the studies where the change of interpretation was induced experimentally), then such an effect can truly be referred to as a retrieval-interference factor that is operative in recognition memory. As argued previously, this factor may or may not be operative in ordinary recognition memory, though it clearly was operative in the abnormal compound-stimulus recognition-memory tasks employed in the previously mentioned studies.

If such a retrieval interference factor is operative in ordinary recognition memory with a constant probability as a function of delay interval, then the effect would probably be to reduce strength recognition retention functions by a constant factor at all retention intervals. Although such retrieval interference effects are of some slight interest in themselves, they are of no significance for the function of recognition tests in studying losses in storage. What would be of greater consequence for the proposition that recognition memory discriminability is linearly related to the strength of the memory trace in storage would be a demonstration that this or some other type of retrieval interference increased systematically in magnitude as a function of retention interval and accounted for all or a substantial part of the forgetting observed on such recognition memory tasks. If this could be demonstrated, it would revive the old uncertainty concerning whether there are any losses in memory storage or whether, in fact, all "forgetting" is simply due to increasing retrieval interference.

I think there is no reason at present to believe that changes in the interpretation of the stimulus item as a function of delay accounts for any substantial portion of the forgetting observed in memory for an AB paired associate, for example. If the interpretation of an item were substantially changing as a function of retention interval, then any previously strongly learned AC associates should be equally seriously affected, as well as more recently and more weakly learned AB associates. Because it is relatively easy to: (1) learn a strong AC paired associate, (2) learn an AB associate to some low criterion, and (3) observe substantial forgetting of the AB associate along with very little forgetting with the AC associate, this specific form of the retrieval interference hypothesis must be rejected.

Another hypothesis is that the A stimulus item is perturbed to some new form by presentation of the AB pair and then systematically drifts back over the retention interval to the interpretation appropriate for old AC pairs. However, any explicit form of this theory would undoubtedly require that the rate of return to the original interpretation to be a function of the strength of the AC paired associate(s). This would yield a prediction that forgetting of an AB pair on a recognition test should be faster following AC learning than following CD learning. The results of Wickelgren (in press) falsifies this prediction.

In conclusion, I think it is possible to say that changes in the interpretation of the stimulus item from learning to test will depress recognition-memory performance. It is not possible to say whether this occurs to any substantial extent under ordinary recognition learning and test conditions, and even if it does occur, it seems unlikely to be responsible in any significant way for the forgetting observed as a function of retention interval in recognition tests. At most, there may be a constant multiplicative retrieval-interference effect on the strength-retention function across all delays.

REFERENCES

Baddeley, A. D., & Ecob, J. R. Reaction time and short-term memory: A trace strength alternative to the high-speed exhaustive scanning hypothesis. *Center for Human Information Processing*, Report No. 13. University of California, San Diego, 1970.

Birnbaum, I. M. General and specific components of retroactive inhibition in the AB, AC paradigm. *Journal of Experimental Psychology*, 1972, **93**, 188–192.

Brown, J. Reciprocal facilitation and impairment of free recall. *Psychonomic Science*, 1968, **10**, 41–42.

Burrows, D., & Okada, R. Serial position effects in high-speed memory search. *Perception and Psychophysics*, 1971, **10**, 305–308.

Ceraso, J., Bader, L., & Silverstein, M. Response latency as a function of list length in paired-associate learning. *Psychonomic Science*, 1970, **19**, 239–240.

Cofer, C. N., Faile, N. F., & Horton, D. L. Retroactive inhibition following reinstatement or maintenance of first-list responses by means of free recall. *Journal of Experimental Psychology*, 1971, **90**, 197–205.

Collins, A. M., & Quillian, M. R. Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 432–438.

Corballis, M. C. Serial order in recognition and recall. *Journal of Experimental Psychology*, 1967, **74**, 99–105.

Corballis, M. C., Kirby, J., & Miller, A. Access to elements of a memorized list. *Journal of Experimental Psychology*, 1972, **94**, 185–190.

DaPolito, F., Barker, D., & Wiant, J. Context in semantic information retrieval. *Psychonomic Science*, 1971, **24**, 180–182.

Davis, J. C., Lockhart, R. S., & Thomson, D. M. Repetition and context effects in recognition memory. *Journal of Experimental Psychology*, 1972, **92**, 96–102.

DeRosa, D. V., & Baumgarte, R. Probe digit recall of items from temporally organized memory sets. *Journal of Experimental Psychology*, 1971, **91**, 154–158.

Egeth, H., Marcus, N., & Bevan, W. Target-set and response-set interaction: Implications for models of human information processing. *Science*, 1972, **176**, 1447–1448.

Eimas, P. D. Subjective reinforcement in the paired-associate learning of retarded and normal children. *Canadian Journal of Psychology*, 1964, **18**, 183–196.

Eimas, P. D., & Zeaman, D. Response speed changes in an Estes' paired-associate "miniature" experiment. *Journal of Verbal Learning and Verbal Behavior*, 1963, **1**, 384–388.

Fischler, I., & Juola, J. F. Effects of repeated tests on recognition time for information in long-term memory. *Journal of Experimental Psychology*, 1971, **91**, 54–58.

Freedman, J. L., & Loftus, E. F. Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 107–115.

Hall, R. F. The effect of number of response alternatives on response frequency and latency in paired-associate learning. *British Journal of Mathematical and Statistical Psychology*, 1969, **22**, 115–130.

Hintzman, D. L. Recognition time: Effects of recency, frequency, and the spacing of repetitions. *Journal of Experimental Psychology*, 1969, **1**, 192–194.

Juola, J. F., & Atkinson, R. C. Memory scanning for words versus categories. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 522–527.

Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. Recognition time for information stored in long-term memory. *Perception and Psychophysics*, 1971, **10**, 8–14.

Karchmer, M. A., & Winograd, E. Effects of studying a subset of familiar items on recall of the remaining items: The John Brown effect. *Psychonomic Science*, 1971, **25**, 224–225.

Kennedy, A. Response latency in the serial learning of short lists. *British Journal of Psychology*, 1968, **59**, 1–5.

Kirsner, K., & Craik, F. I. M. Naming and decision processes in short-term recognition memory. *Journal of Experimental Psychology*, 1971, **88**, 149–157.

Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and recognition time. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 291–295.

Lappin, J. S., & Disch, K. The latency operating characteristic: I. Effects of stimulus probability on choice reaction time. *Journal of Experimental Psychology*, 1972, **92**, 419–427. (a)

Lappin, J. S., & Disch, K. The latency operating characteristic: II. Effects of visual stimulus intensity on choice reaction time. *Journal of Experimental Psychology*, 1972, **93**, 367–372. (b)

Light, L. L., & Carter-Sobell, L. Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 1–11.

Lively, B. L. Speed/accuracy tradeoff and practice as determinants of stage durations in a memory-search task. *Journal of Experimental Psychology*, 1972, **96**, 97–103.

McGeoch, J. A., & Irion, A. L. *The Psychology of Human Learning*. New York: Longmans, Green, 1952.

Metlay, W., Handley, A., & Kaplan, I. T. Memory search through categories of varying size. *Journal of Experimental Psychology*, 1971, **91**, 215–219.

Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, **1**, 242–300.

Millward, R. Latency in a modified paired-associative learning experiment. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 309–316.

Morin, R. E., DeRosa, D. V., & Stultz, V. Recognition memory and reaction time. *Acta Psychologica*, 1967, **27**, 298–305.

Morin, R. E., DeRosa, D. V., & Ulm, R. Short-term recognition memory for spatially isolated items. *Psychonomic Science*, 1967, **9**, 617–618.

Murdock, B. B., Jr. Response latencies in short-term memory. *Quarterly Journal of Experimental Psychology*, 1968, **20**, 79–82.

Norman, D. A. Acquisition and retention in short-term memory. *Journal of Experimental Psychology*, 1966, **72**, 369–381.

Norman, D. A., & Wickelgren, W. A. Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology*, 1969, **6**, 192–208.

Okada, R. Decision latencies in short-term recognition memory. *Journal of Experimental Fsychology*, 1971, **90**, 27–32.

Pachella, R. G., & Fisher, D. Hick's law and the speed-accuracy tradeoff in absolute judgment. *Journal of Experimental Psychology*, 1972, **92**, 378–384.

Peterson, L. R. Paired-associate latencies after the last error. *Psychonomic Science*, 1965, **2**, 167–168.

Postman, L., Stark, K., & Fraser, J. Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 672–694.

Rabbitt, P. M. A., & Vyas, S. M. An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologica*, 1970, **33**, 56–76.

Reed, A. Speed-accuracy tradeoff in recognition memory. *Science*, 1973, **18**, 574–576.

Schlag-Rey, M., Groen, G., & Suppes, P. Latencies on last error in paired associate learning. *Psychonomic Science*, 1965, **2**, 15–16.

Schouten, J. F., & Bekker, J. A. M. Reaction time and accuracy. *Acta Psychologica*, 1967, **27**, 143–153.

Sternberg, S. High-speed scanning in human memory. *Science*, 1966, **153**, 652–654.

Sternberg, S. Retrieval of contextual information from memory. *Psychonomic Science*, 1967, **8**, 55–56.

Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, **57**, 421–457.

Suppes, P., Groen, G., & Schlag-Rey, M. A model for response latency in paired-associate learning. *Journal of Mathematical Psychology*, 1966, **3**, 99–128.

Swanson, J. M., & Briggs, G. E. Information processing as a function of speed vs. accuracy. *Journal of Experimental Psychology*, 1969, **81**, 223–229.

Swensson, R. G. The elusive tradeoff: Speed vs. accuracy in visual discrimination tasks. *Perception and Psychophysics*, 1972, **12**, 16–32.

Swensson, R. G., & Edwards, W. Response strategies in a two-choice reaction task with a continuous cost for time. *Journal of Experimental Psychology*, 1971, **88**, 67–81.

Theios, J. Prediction of paired-associate latencies after the last error by an all-or-none learning model. *Psychonomic Science*, 1965, **2**, 311–312.

Thomson, D. M. Context effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 497–511.

Tulving, E., & Thomson, D. M. Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology*, 1971, **87**, 116–124.

Waugh, N. C. Retrieval time in short-term memory. *British Journal of Psychology*, 1970, **61**, 1–12.

Wickelgren, W. A. Exponential decay and independence from irrelevant associations in short-term recognition memory for serial order. *Journal of Experimental Psychology*, 1967, **73**, 165–171.

Wickelgren, W. A. Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology*, 1972, **9**, 418–455.

Wickelgren, W. A. Single-trace fragility theory of memory dynamics. *Memory and Cognition*, in press.

Wickelgren, W. A., & Norman, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, **3**, 316–347.

Wiggins, J. S. Two determinants of associative reaction time. *Journal of Experimental Psychology*, 1957, **54**, 144–147.

Wilkins, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 382–385.

Winograd, E., & Conn, C. P. Evidence from recognition memory for specific encoding of unmodified homographs. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 702–706.

Winograd, E., Karchmer, M. A., & Russell, I. S. Role of encoding unitization in cued recognition memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 199–206.

Yellott, J. I., Jr. Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*, 1971, **8**, 159–199.