

Retention in perceptual memory: A review of models and data

DONALD LAMING and PETER SCHEIWILLER
University of Cambridge, Cambridge, England

Five contemporary models are surveyed and four experiments reanalyzed in a review of present knowledge of how the discriminability of two tones or of two lines of different lengths varies with the temporal interval (1 to 60 sec) between their presentations. Although discriminability (d') decreases roughly in proportion to the logarithm of the time elapsed, existing experiments have difficulty in discriminating between the alternative models, and there are practical constraints on repeating those experiments with significant improvements in precision. However, previous work suggests that all the models might fail at very short intervals (< 1 sec); certain experiments with auditory stimuli have found d' to decrease at these short intervals, reversing its principal trend. A fifth experiment attempts to replicate this finding with visual stimuli; one subject, but not the other, shows a clear maximum in discriminability at an interstimulus interval of 0.25 sec. This finding is compared, speculatively, with reminiscence and with specific psychophysical sensitivity to changes in the stimulus field.

If two stimuli—two tones or two lengths of line—are presented in succession for comparison, the accuracy of that comparison generally decreases with increase in the temporal interval between the stimuli. We are concerned here with the quantitative nature of the relation between the accuracy of the comparison and the time elapsed; and we shall measure that accuracy by a threshold value [the stimulus difference that affords 75% correct two-alternative forced-choice (2AFC) responses] or by the discriminability index d' for a fixed stimulus difference.

This topic has a long history, but has been largely neglected in this present century, possibly because of the extreme difficulty of obtaining conclusive results, as will appear. Nevertheless, we think it worth attention for these three reasons:

(1.) It is easier and much more common to study forgetting in the short term using words or digits or non-sense syllables, but such verbal stimuli suffer a unique disadvantage. There are only a finite (or at most a countable) number of such stimuli, which are, of necessity, known to the subject and may be envisaged as discrete neighborhoods in a multidimensional space of acoustic waveforms. Now, these stimuli (indeed, any stimulus) may be recalled spontaneously by the subject; and spontaneous recall, being a conscious mental experience, creates its own trace in memory which may itself be recalled in due turn. If the stimuli are discrete, then recall, provided it is not too long delayed, is virtually error-free; that

is, knowledge of the discrete set of possible stimuli may be invoked to correct any potential errors in recall. And, by rehearsing the stimuli in this way, it is possible to retain a limited amount (e.g., the digit span) of verbal material without error for an arbitrary period of time; this is very well known. But tones and lengths of line may take a *continuum* of possible values; and, although such stimuli may be recalled spontaneously, it is, for this reason, not possible to correct errors in recall.

In experiments with verbal stimuli, it is necessary to impose a distractor task to prevent rehearsal (e.g., the Brown-Peterson paradigm); but without collateral observations one cannot know how effective such a distractor task is, nor whether it introduces explicit interference with the material to be remembered. Using a continuum of stimuli, however, renders this device unnecessary.

(2.) One might expect the retention of any stimulus to depend on its psychological properties—words have meanings and associations and these have long been known to be related to memorability (Noble, 1952a, 1952b). In comparison, the stimuli employed in the experiments below are much simpler and their psychological properties better understood (see Laming, in press). In addition, signal-detection theory provides a well-tried model for the analysis of the data.

(3.) There are a few experiments in which discriminability has, paradoxically, improved with lapse of time, up to a second or so, and it is possible that further study of the conditions under which this is observed might promote some radical development of our understanding of retention.

The study of the retention of simple stimuli in perceptual memory has a long history, which begins with a variety of casual observations about mnemonic function reported by Weber (1834/1978, 1846/1978; see Ross &

We thank A. D. Baddeley and D. W. Massaro for their comments on a previous draft of this paper. The two original experiments reported here were carried out by the second author as part of his third-year undergraduate study at the University of Cambridge.

The authors' mailing address is: Department of Experimental Psychology, University of Cambridge, Cambridge, England CB2 3EB.

Murray, 1978) in connection with his measurements of sensory thresholds. Inspired by Weber's observations, Hegelmaier (1852) published what is possibly the first psychological experiment (using "experiment" in its modern sense of a systematic, planned, *body* of observations) under the title "On Memory for the Length of a Line." Hegelmaier's work contains some very elementary errors of procedure, but is nevertheless of great historical significance, since it is the source of the method of constant stimuli. Hegelmaier's experimental paradigm was subsequently used for the study of memory by others, especially Wolfe (1886) and Lewy (1895), whose data we reanalyze below. These two studies, and many others of lesser precision, were reviewed by Kennedy (1898), whose article, so far as we can discover, is the most recent survey of this small backwater of memory research.

We began by replicating Hegelmaier's experiment with the benefit of modern instrumentation, modern experimental design, and modern statistical analysis. Add to this a handful of studies published around 1970, conducted, it would seem, without awareness of the 19th century work, and the time is ripe for a further review. In this review, we list the alternative models that have been proposed and then show, by reanalysis, the extent to which the best of the extant sets of data discriminate between them. This exercise clears the decks prior to further research in this area.

FIVE MODELS FOR THE RETENTION OF SIMPLE STIMULI

If we were proposing a *theory* of perceptual memory, we would need, as a very difficult preliminary, to justify a specific measure of memory "strength." But, instead, we look on the models that follow simply as calculational vehicles for the analysis of experimental data and resolve this question pragmatically. It happens that all the psychophysical properties of a discrimination between two separate stimuli are nicely scaled by the value of the threshold (Laming, in press, chap. 5); so we shall use the inverse threshold or, what is proportionate, the discriminability index d' as a pragmatic measure of the strength of memory.

The Exponential Model

The exponential model is

$$d'(t) = d'(0) e^{-at}, \quad (1)$$

where t is the retention interval. This equation makes forgetting analogous to the loss of charge from a capacitor, or the decay of radioactive material, or the damping of the oscillations of a pendulum. This idea seems to have been first proposed by Nipher (1876; see Stigler, 1978), who used the third of these analogies.

The Logarithmic Model

Wolfe (1886) varied the retention interval from 1 to 60 sec and fitted the purely empirical equation

$$(P.\text{correct})/(P.\text{wrong}) = k/(\log t) + c, \quad (2)$$

which is very similar to the equation fitted by Ebbinghaus (1885/1913/1964; see the translation by Ruger and Busenius, p. 77) to his own data on the retention of lists of nonsense syllables over much longer periods of time (19 min to 31 days). The recall of a variety of different kinds of material seems to depend on retention interval via the medium of its *logarithm* (i.e., $\log t$ in Equation 2; see Woodworth & Schlosberg, 1955, pp. 724-728), and as a representative of this kind of relation we explore the more modern equation

$$d'(t) = d'(1) (1 - a \ln t). \quad (3)$$

The Diffusion Model

Suppose the presentation of the standard stimulus to be represented by a random sample from a normal distribution, as in signal-detection theory. Suppose that during the retention interval this representation suffers many small independent random perturbations occurring at a uniform rate, and so evolves in a manner analogous to Brownian motion. The cumulative effect of these perturbations can be represented by an increase in the variance of the normal distribution, an increase proportional to the time elapsed. The presentation of the comparison stimulus is represented by a random sample from the other distribution in the signal-detection model; and the combined variance of the discrimination between the two stimuli increases with the retention interval as $(1 + at)$, so that

$$d'(t) = d'(0) (1 + at)^{-1/2}. \quad (4)$$

Kinchla and Smyzer (1967) first proposed this model and reported experiments on the discrimination of visual position and of auditory amplitude in good agreement with Equation 4, but explored only the range of retention intervals from 0.5 to 2 sec. Kinchla and Allan (1969) developed this idea into a model of visual movement perception, and Keller and Kinchla (1968) and Kinchla and Allan (1970) contributed further experiments, but still only over the range 0.5 to 2 sec.

The Dual-Trace Model

Wickelgren (1969) substantially replicated Wolfe's experiment with the additional insertion of an interfering tone during the retention interval. He fitted his estimates of sensitivity for retention intervals ranging from 1 to 180 sec to the equation

$$d'(t) = Ae^{-at} + Be^{-bt}. \quad (5)$$

Estimates of a from three subjects ranged from 0.115 to 0.29 sec^{-1} (corresponding to a short-term trace) and of b from 0 to 0.025 sec^{-1} (the long-term trace).

The Storage-Interference Model

Massaro (1970a) proposed a storage-interference model specifically for the memory of the frequency of a tone.

Memory strength was assumed to increase according to a negative exponential function during the presentation of the standard tone and also during any blank interval (in practice, of 0.5 sec duration) that followed the standard; and to decrease during the presentation of an interfering tone according to the equation

$$d'(t) = d'(0) \exp[-a(1 - e^{-bt})]. \quad (6)$$

Since the memory of a tone deteriorates during blank intervals much longer than 0.5 sec, Equation 6 might plausibly describe this forgetting.

FOUR EXPERIMENTS

We reanalyze below the data from three experiments by Lewy (1895), Wickelgren (1969), and Wolfe (1886), and add a fourth of our own. The numerical data for each of the three experiments have been published and cover a wide range of intervals (at least 1 to 60 sec); we believe they are the only three. In addition, since the original publications by Wolfe and Lewy are not readily accessible, we republish the best of their data in Figures 1 and 2.

Wolfe (1886)

Wolfe used a set of about 300 metal reeds to produce tones in the frequency range 32 to 1024 Hz. In his experiment, he presented a standard tone of 1 sec duration, followed by a silent interval in the range 1 to 60 sec and then a comparison tone which might be equal to, or greater or less (by 4 Hz) than, the standard. Figure 1 displays the proportions of correct judgments from the two of Wolfe's subjects who contributed the greatest numbers of observations. Each data point is based on 297 (Lehmann) or 841 (Wolfe) separate trials. The standard fre-

quency assumed one of 11 different values in the range 144 to 1004 Hz and was, in addition, varied by a further ± 4 Hz to preclude absolute judgment of the comparison tone.

Wickelgren (1969)

Wickelgren substantially replicated the work of Wolfe, with the additional interposition of an interfering tone between the standard and the comparison. The standard tone (ranging from 400 to 490 Hz in 10-Hz steps) lasted 3 sec. The interfering tone was always 930 Hz and filled the entire retention interval, which ranged from 1 to 180 sec. The comparison tone was either equal to the standard (on half the trials) or differed from it by ± 10 Hz. Subjects expressed their judgments on a 9-point scale, from which estimates of d' were calculated. The numerical estimates are included in the legend of Figure 6 in Wickelgren (1969, p. 26), and the d' scores are plotted in Figure 8 (p. 32) of the same article.

Lewy (1895)

As his stimulus to be remembered, Lewy used the distance between two small ivory disks (1.5-mm diam) viewed against a black cloth background from a distance of 50 cm. One disk was fixed; the other could be moved on a black silk thread and its position read off from a millimeter scale. On each trial, the subject viewed, for 5 sec, a standard separation between the two disks (which ranged on different trials from 60 to 200 mm). During the retention interval, the subject closed his eyes and the experimenter displaced the movable disk. At the end of the retention interval, which ranged from 1 to 60 sec, the subject reopened his eyes and gave directions to the experimenter to restore the movable disk, as near as he (the subject) could judge, to its previous position—a kind of method of average error. If the standard deviation of the variable error of adjustment is expressed as a percentage of the standard stimulus and averaged over the different standards, then Figure 2 shows the inverses of those averages—the inverses being proportional to d' scores—for two subjects. Each data point in Figure 2 is based on 96 separate adjustments.

Our Experiment 1

To these three, we add our replication of Hegelmaier's experiment.

Our stimuli were generated in reverse on a video display unit, appearing as horizontal black lines on a bright green ground. The ground extended 23 cm horizontally and 13.5 cm vertically, but only the central 16×9.5 cm was used for the stimuli. On each trial, a standard line was presented for 0.5 sec in some randomly chosen location within the central part of the ground. The standard was chosen to have one of seven different lengths at random, ranging from 3° to 12° of visual angle. After a retention interval between 1 and 64 sec, a horizontal comparison line was presented, also for 0.5 sec, in some independent random location. The comparison line was

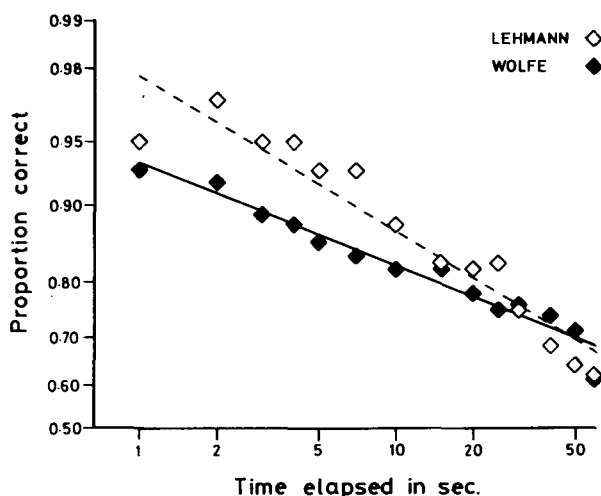


Figure 1. Proportions of correct responses from two subjects judging a comparison tone that differed by ± 4 Hz from, or was equal to, the standard. The straight lines are least squares regression lines, and the data are taken from Wolfe (1886, p. 555, Tabelle VIII).

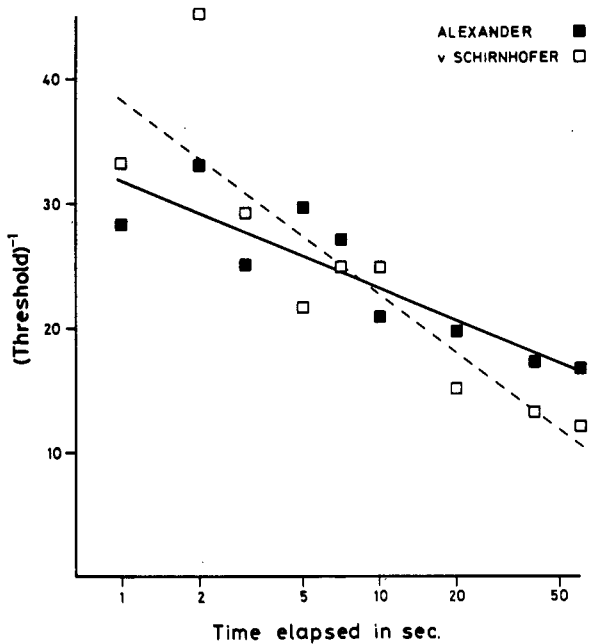


Figure 2. Inverse probable errors of adjustment for two subjects attempting to recreate a previously seen separation between two small white disks together with least squares regression lines. The data are taken from Lewy (1895, p. 242, Tabelle I).

either shorter or longer than the standard by some multiple of 2%. The presentations of the standard and of the comparison lines were each preceded by a brief auditory warning signal, and each response was followed by knowledge of results before the next trial commenced.

The percentage difference in length between comparison and standard was initially set at 10% and thereafter adjusted up or down according to a staircase procedure which converged on that difference in length, which afforded 71% correct responses (see Levitt, 1971). A separate staircase was generated for each different retention interval, so that the experiment consisted of seven subexperiments randomly interleaved; but within such a subexperiment no distinction was made between the different lengths of the standard line. The two authors served as subjects, and each completed five experimental sessions. Each session continued until 11 reversals of the direction of adjustment had been generated in each of the seven staircases; this took about 100 min. Estimates of the inverse thresholds for each subject in this experiment are shown in Figure 3.

Other Experimental Work

There have, of course, been other experiments on the retention of simple stimuli of lesser precision or covering a lesser range of retention intervals. The 19th century work is catalogued by Kennedy (1898). Since that time there have been, among others, experiments by Bachem (1954) on the retention of a pure tone (this experiment is discussed below); by Adams and Dijkstra

(1966) on the accuracy of reproduction of a movement; by Posner and Konick (1966) on the recall of the location of a point on a line, and a replication by Dale (1973) on the recall of the position of a dot within a square; and finally by Phillips (1974) on the recognition of matrices of black and white squares. The experiments reanalyzed here were selected as being the most informative with respect to the question of present interest.

ANALYSIS OF THE DATA

Least Squares Analysis

It is generally found that the standard deviation of independent estimates of a threshold is itself proportional to the threshold (Crozier, 1936), and there are theoretical reasons why this should be so (Laming, in press, chap. 5). Accordingly, we minimized the sums of squared deviations of \ln threshold or $\ln d'$ about the predictions of each of our five models. Table 1 shows the minimum mean square deviation for each model and each separate subject in the four experiments; and it is immediately apparent that all the models represent the data to about the same degree of precision. The magnitude of each statistic reflects the concordance between model and data compounded with the precision of the individual experiment, and for this latter reason comparisons between different experiments, or even between different subjects within the one experiment, may not be meaningful. But comparisons within each row (i.e., between models) are directly meaningful, since mean squared deviation is the statistic that most accurately reflects the concordance between model and data. And we immediately remark that Wickelgren's experiment shows a profile of squared deviations very different from the other three.

It is not possible to evaluate statistical comparisons between the different models, except that the exponential model is a special case of both the dual-trace model (deleting the long-term trace in Equation 5) and of the storage-

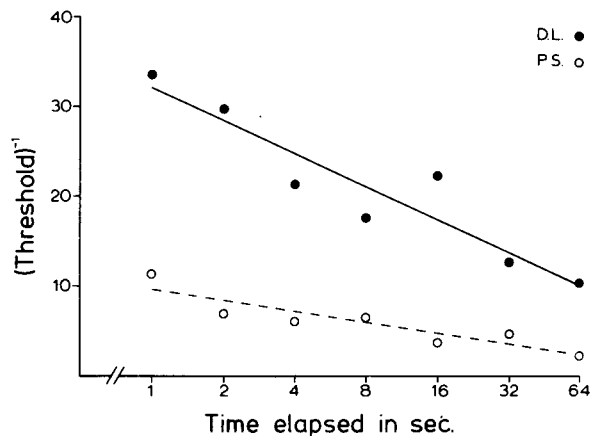


Figure 3. Inverse thresholds from Experiment 1, together with least squares regression lines.

Table 1
Minimum Mean Square Deviations for Each Subject in the Four Experiments

Subject	Exponential Model	Logarithmic Model	Diffusion Model	Dual-Trace Model	Storage-Interference
Wolfe, 1886					
A.L.	0.008	0.020	0.050	0.009	0.008
H.K.W.	0.019	0.022	0.026	0.023	0.021
Wickelgren, 1969					
B.F.	0.672	0.398	0.161	0.015	0.031
D.W.	0.209	0.108	0.095	0.005	0.005
M.S.	0.327	0.244	0.263	0.044	0.028
Lewy, 1895					
Al	0.018	0.012	0.012	0.013	0.011
v. Sch.	0.060	0.031	0.030	0.037	0.031
Present Experiment 1					
D.L.	0.050	0.025	0.031	0.031	0.042
P.S.	0.076	0.048	0.056	0.049	0.083

interference model (letting $b \rightarrow 0$ while ab remains fixed in Equation 6). So the reduction in the total sum of squared deviations occasioned by the introduction of the additional two (or one) free parameters permits the calculation of an F ratio. In Wickelgren's experiment, the dual-trace and the storage-interference models achieve a significant improvement with respect to the exponential model, significant in every case to at least 0.01. In the other three experiments, only one such comparison is significant [in Lewy's experiment the storage-interference model shows a significant improvement over the exponential model in the analysis of von Sch.'s data; $F(1,6) = 7.671, p < .05$]; and, much more to the point, the technical reductions in the total sum of squares when fitting both the dual-trace and the storage-interference models to H.K.W.'s data (Wolfe, 1886) were so small that they were lost in the rounding errors of the calculations. This means that in this one case the F ratios were significantly *smaller* than one would expect by chance.

The reasons for this difference become apparent on examining the least squares estimates of the model parameters. In Wickelgren's experiment, the rate of decay of the long-term trace (parameter b in Equation 5) was estimated (by our weighted least squares procedure) to be 0.000, 0.010, and 0.000 for the three subjects, respectively, so that within the range of retention intervals explored d' tended to a stable limiting value which was estimated to be 0.30, 0.18, and 0.13, respectively. The single exponential model, on the other hand, requires this limiting value to be zero, and the value estimated from H.K.W.'s data (when fitting the dual-trace model) was indeed negligible ($< 5 \times 10^{-5}$). We think the difference is due to Wickelgren's use of an interfering tone to fill the retention interval. Massaro (1970b) has shown that, with short retention intervals (≤ 4 sec), the interfering tone makes the discrimination of frequency more difficult. But over much longer intervals, the interfering tone might function as a point of reference serving to preserve the memory at a low level, somewhat analogous to a brief

note scribbled on a memorandum pad. Looking back at the first of the arguments in our introduction, we recommend that future work on the question of the discrimination of auditory frequency, at the least, should use no interfering stimulus.

Goodness-of-Fit

It is not possible to assign any significance level to individual statistics in Table 1 because there is no way of calculating what those mean squared deviations should be a priori. But such a calculation is possible with our own data because we have the individual judgments at hand. We model each response as an independent binomial trial with a probability that varies as a normal integral with respect to the stimulus difference; and the inverse standard deviation of the normal integral is made to vary with the retention interval according to one or another of our five models. On this basis, we calculated a likelihood-ratio statistic for all except the dual-trace model (where technical difficulties supervened), and the results are displayed in Table 2.

We emphasize at once that all of the models account for the data to within the limits of experimental error. The values of the chi-square statistics are close to their respective numbers of degrees of freedom, which are also their expected values. But there is, as above, a valid comparison between the storage-interference model and the single exponential. The storage-interference model demonstrates a significant improvement in fit for Subject D.L. ($\chi^2 = 5.618, p < .05$) but not for Subject P.S.

The statistics in Table 2 illustrate the difficulty of research into the retention of simple stimuli. The numerical predictions (after the optimal adjustment of at least two free parameters) from the models so far proposed are very close to each other in relation to the precision that can be achieved in any one experiment. To improve this precision one might:

- (1) Record a much larger number of observations from each subject—though that will *not* resolve a choice between models, all of which are only approximations to the true state of nature.
- (2) Extend the range of retention intervals compared. For example, extrapolation of the logarithmic model (Equation 3) indicates that $d'(t)$ will reach zero when $t = \exp(1/a)$, and least squares estimates of the param-

Table 2
The Concordance of Each Model with the Data from Experiment 1

Model	Subject D.L.		Subject P.S.	
	χ^2	df	χ^2	df
Single Exponential (Equation 1)	55.183	46	76.575	71
Logarithmic (Equation 3)	47.660	47	72.858	71
Diffusion (Equation 4)	49.688	47	73.868	71
Storage-Interference (Equation 6)	49.565	46	74.461	70

eter a indicate that this critical duration varies in the experiments here from 2 to 70 min. We would not, however, expect the memory to disappear entirely within that finite time.

There are obvious practical difficulties in collecting a sufficient number of judgments combined with so long a retention interval. Nevertheless, Bachem (1954) has reported such an experiment, studying the accuracy of discrimination of auditory frequency when the standard and comparison tones were separated by, variously, 1, 3, 15, and 60 sec, 1 h, 1 day, and 1 week. For two subjects without the faculty of "absolute pitch," Bachem's results appear to conform (very approximately) to the logarithmic model up to, perhaps, 1 h; but thereafter the rate of loss of discriminability is much less than extrapolation of Equation 3 would indicate. Two other subjects possessing "absolute pitch" reached an asymptotic lower limit of discriminability after about 1 min. This lower limit is mediated by categorizing the chroma of the tone relative to the notes of the musical scale and remembering that verbal label; it works only for frequencies within the usual musical range (up to 4 kHz).

(3) Look for some related phenomenon which might indicate how to increase the *variety* of relevant observations. Such a phenomenon is, we believe, the dramatic loss of discriminative efficiency reported in some experiments with very short interstimulus intervals, and we now turn to an examination of this phenomenon.

NONMONOTONIC RETENTION FUNCTIONS

All of our five models describe a strictly monotonic loss of discriminative efficiency as the retention interval increases. So, if it could be established that under certain conditions discrimination actually *improved* with increase in the retention interval, a radical revision of all five models would be required.

The data from Lewy in Figure 2 and from one of Wolfe's subjects in Figure 1 show a lower threshold or greater proportion of correct responses after an interval of 2 sec than after 1 sec. Although this finding is counterintuitive, it is possibly empirical confirmation of Weber's (1834/1978) report (see Ross & Murray, 1978, pp. 94-107) that successively presented stimuli are more easily discriminated than simultaneous ones. More recently, Tanner (1961, Figure 9) has shown that the efficiency of discriminating two brief 1-kHz tones of different amplitudes presented successively in Gaussian noise varies in a remarkable way with the interval between their presentations. Efficiency was maximal at an interstimulus interval of 600 msec. At longer intervals, it decreased roughly according to the diffusion model (Equation 4); at shorter intervals (300, 100, and 50 msec), it decreased dramatically. This result has been replicated for monaural discriminations by Sorkin (1966), who also showed that a dichotic discrimination behaved differently. When the two tones (of 50 msec duration) were presented to different ears, peak efficiency was obtained at an interstimulus

interval of about 200 msec, with a less dramatic decrease at shorter intervals. The decrease of efficiency over these shorter intervals has been studied in detail by Taylor and Smith (1975), who report that the stimulus needs to be brief to demonstrate the effect reliably.

Since all this recent work has used auditory discriminations, we conducted a second experiment with our visual task, exploring a range of shorter intervals of time.

Our Experiment 2

We used the same stimuli presented exactly as in Experiment 1 and for the same duration (0.5 sec). The sequence of events on each trial was as before, except that the difference between the standard and comparison lines was now fixed and the subject expressed his posterior confidence, on a 6-point scale, that the comparison line was shorter/longer than the standard, using the paradigm introduced by Swets, Tanner, and Birdsall (1961, pp. 325-329). Seven different interstimulus intervals, measured between the offset of the standard line and the onset of the comparison, were compared: nominally 0 (i.e., immediate succession), 1/8, 1/4, 1/2, 1, 2, and 4 sec. In addition, a true simultaneous condition was included in which both lines appeared together, but in different random locations. These eight experimental conditions were randomly interleaved in the sequence of experimental trials.

The two authors again served as subjects, completing a total of 500 trials under each experimental condition. In the light of the results of Experiment 1, the difference in length between the standard and comparison lines was fixed at 4% for D.L. and 8% for P.S.

Results. The 500 trials under each different interstimulus interval were analyzed according to the normal, equal variance, signal-detection model, and the maximum-likelihood estimates of *d'* are displayed in Figure 4. Of the 16 goodness-of-fit statistics, 4 were significant at .05 (but not at .01). The corresponding sets of data points

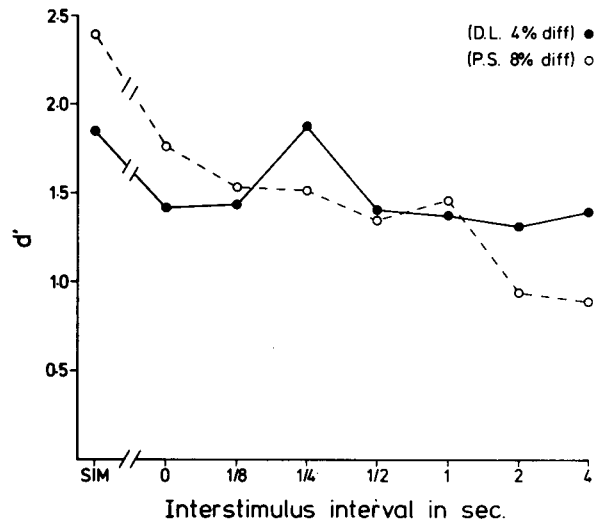


Figure 4. Estimates of *d'* for each interstimulus interval in Experiment 2.

deviated from the normal operating characteristic in showing too sharp a "knee"; this deviation is exemplified by the data for the simultaneous, 1/4-, and 1/2-sec conditions in Figure 5.

We next compared the estimates of d' from the different experimental conditions and discovered that those conditions could be partitioned (differently for each subject), as set out in Table 3. Estimates from conditions within the same group do not differ significantly; the differences between the common d' estimates from different groups are highly significant. Thus, Subject D.L. (but not P.S.) shows a peak in discriminative sensitivity at an interstimulus interval of 1/4 sec. This appears to replicate the findings of Tanner (1961) and Sorkin (1966); but we do not know why only one subject showed this effect. It is not the case that the significance of D.L.'s data is only marginal; the probability of obtaining a χ^2 as great as 26.358 by chance is less than 10^{-6} .

Finally, we looked for shifts in criteria within those groups of conditions in Table 3 which admit a common estimate of d' . Subject P.S. showed no such changes. But D.L. showed a systematic change as the interstimulus interval increased through the values 1, 2, and 4 sec, whereby the criteria assumed lower values, leading to an increased proportion of "longer" judgments. Since the temporal difference between intervals of 1 and 2 and 4 sec is readily perceptible, it is plausible that the subject should adjust his criteria in this way. And the nature of the adjustment accords with a great deal of earlier work summarized in Köhler's (1923) model of a fading memory trace, a model with which he sought to explain the preponderance of negative time-errors in psychophysical dis-

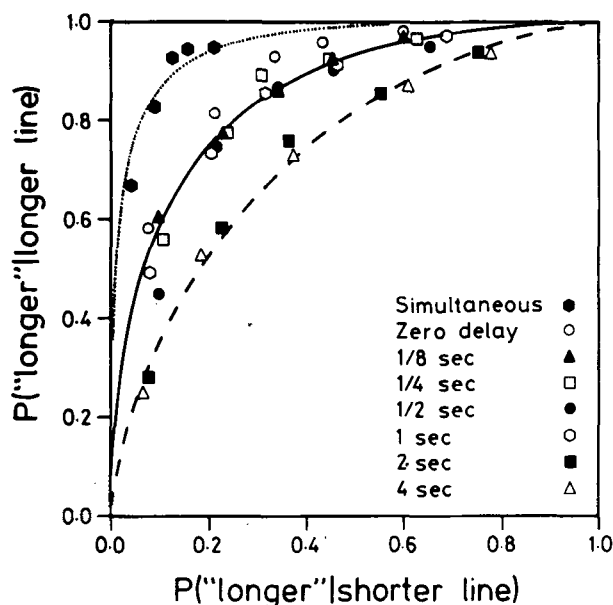


Figure 5. A signal-detection plot of the proportions of each kind of judgment made by Subject P.S. at each interstimulus interval in Experiment 2. Three operating characteristics are shown for comparison, with d' equal to 0.916, 1.517, and 2.393.

Table 3
Comparisons of d' Estimates from Experiment 2

Interstimulus Intervals	Mean d'	χ^2	df
Subject D.L.			
Sim, 1/4	1.865	0.031	1
0, 1/8, 1/2, 1, 2, 4	1.390	0.769	5
Residual		26.358*	1
Subject P.S.			
Sim	2.393		
0, 1/8, 1/4, 1/2, 1	1.517	7.800	4
2, 4	0.916	0.126	1
Residual		128.898*	2

*Significant at .001.

crimination (see Woodworth & Schlosberg, 1955, pp. 220-228).

DISCUSSION

We now take up certain more general questions posed by our analysis.

Is Perceptual Memory MEMORY?

That is, is the faculty we have been studying the same as, or related to, that studied by many others with nonsense syllables, digits, words, sentences, pictures, and so on? We think the answer should be "Yes," because the trend of memory loss in the other experiments we have analyzed can be approximately described by the same kind of equation (3) that has been found applicable to the forgetting of the other kinds of materials listed above, with retention tested over much longer intervals of time (see Woodworth & Schlosberg, 1955, pp. 724-728). Physical processes that evolve even *approximately* in logarithmic time are, at the least, uncommon; and for this reason we think it likely that the mnemonic function we have been studying here has much in common with those other mnemonic functions studied elsewhere.

This is an important issue, because, if the relation between memory for simple stimuli and memory for words, pictures, etc., is conceded, then the techniques and theoretical groundwork of sensory discrimination and of signal-detection experiments can be brought to bear on the study of memory more acutely than they have been hitherto. Tanner (1961) has already indicated what might be achieved by this means.

What is the Shape of the Forgetting Function?

Frankly, we do not know. When forgetting is examined over a sufficiently wide range of retention intervals (Woodworth & Schlosberg, 1955), it clearly is not an exponential decay; the loss of memory increases more nearly in proportion to the *logarithm* of the time elapsed. The logarithmic relation cannot be mathematically correct—it assumes negative values after a finite lapse of time and tends to infinity as the retention interval tends to zero. However, it fits most of the data here (Wickelgren's experiment excepted).

What of the Peak in Sensitivity for Subject D.L. After a Retention Interval of ¼ sec?

This could be no more than a chance deviation, but the significance level attaching to this result is extreme. Since similar phenomena have been reported by Tanner (1961) and others listed above, we think mere chance is unlikely to be the explanation. Taking this result at its face value, two possible, though speculative, analogies spring to mind:

Reminiscence. Ballard (1913) found that schoolchildren set to learn poetry showed an improvement from an immediate written recall to an unexpected second recall up to 5 days later. Likewise, Ward (1937) demonstrated an improvement in the recall of nonsense syllables, presented for massed learning on a memory drum, for groups of subjects tested after ½, 2, and 5 min relative to a control group tested immediately. There might be a superficial similarity to these two results and others like them, different kinds of material passing through a similar phase of reminiscence followed by forgetting, albeit at very different rates.

The comparison of difference discrimination and increment detection. The discrimination between two separate stimuli of levels X and (X + ΔX) is psychophysically different from the detection of an increment ΔX added to a continuous background of level X, and may even exhibit a lower threshold (see Laming, in press, chaps. 1 and 4, and Leshowitz, Taub, & Raab, 1968, Figures 9 and 10, for explicit examples). It might be meaningful to compare the zero-delay condition in Experiment 2 to an increment detection and the ¼-sec condition to an optimum difference discrimination. And it might eventually turn out that the decay described by our five models is of the same psychological nature as ordinary forgetting, but that the nonmonotonicity at very short interstimulus intervals is of sensory origin.

What Further Experimental Work Would Be Desirable?

We have reanalyzed the best existing data with respect to the temporal course of the retention of sensory discriminability, and that reanalysis has failed to discriminate reliably between any of the different models that have been proposed. Experiments of a sufficiently increased size with much larger numbers of trials and covering much wider ranges of retention interval appear impracticable, but might ultimately prove to be the only way to study this question. We suspect that examining the possible occurrence of a peak in discriminative sensitivity after some short interstimulus interval offers a more hopeful way forward. In preliminary studies, Taylor and Smith (1975, p. 1500) found no variation in sensitivity for long-duration (0.5 sec) auditory signals; such variation was found only for short signals, which is perhaps why our Experiment 2 was relatively uninformative. But if Tanner's peak in discriminative efficiency can be shown to be a robust phenomenon, and if it can be related to our present understanding of sensory discrimination, there might be, in

addition, some profound intuitions about how to model the retention of simple stimuli over longer periods of time.

REFERENCES

- ADAMS, J. A., & DIJKSTRA, S. (1966). Short-term memory for motor responses. *Journal of Experimental Psychology*, **71**, 314-318.
- BACHEM, A. (1954). Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, **26**, 751-753.
- BALLARD, P. B. (1913). Obliviscence and reminiscence. *British Journal of Psychology*, Monograph Supplement.
- CROZIER, W. J. (1936). On the variability of critical illumination for flicker fusion and intensity discrimination. *Journal of General Physiology*, **19**, 503-522.
- DALE, H. C. A. (1973). Short-term memory for visual information. *British Journal of Psychology*, **64**, 1-8.
- EBBINGHAUS, H. (1885). *Über das Gedächtnis*. Leipzig: Duncker & Humblot. [Memory (H. A. Ruger & C. E. Busseensius, Trans.). New York: Teachers College, Columbia University, 1913. Republished New York: Dover, 1964.]
- HEGELMAIER, F. (1852). Über des Gedächtnis für Linearanschauen. *Archiv für Physiologisch Heilkunde*, **11**, 844-853.
- KELLER, W., & KINCHLA, R. A. (1968). Visual movement discrimination. *Perception & Psychophysics*, **3**, 233-236.
- KENNEDY, F. (1898). On the experimental investigation of memory. *Psychological Review*, **5**, 477-499.
- KINCHLA, R. A., & ALLAN, L. G. (1969). A theory of visual movement perception. *Psychological Review*, **76**, 537-558.
- KINCHLA, R. A., & ALLAN, L. G. (1970). Visual movement perception: A comparison of sensitivity to vertical and horizontal movement. *Perception & Psychophysics*, **8**, 399-405.
- KINCHLA, R. A., & SMYZER, F. (1967). A diffusion model of perceptual memory. *Perception & Psychophysics*, **2**, 219-229.
- KÖHLER, W. (1923). Zur Theorie des Sukzessivvergleichs und der Zeitfehler. *Psychologische Forschung*, **4**, 115-175.
- LAMING, D. R. J. (in press). *Sensory analysis*. Academic Press.
- LESHOWITZ, B., TAUB, H. B., & RAAB, D. H. (1968). Visual detection of signals in the presence of continuous and pulsed backgrounds. *Perception & Psychophysics*, **4**, 207-213.
- LEVITT, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, **49**, 467-477.
- LEWY, W. (1895). Experimentelle Untersuchungen über das Gedächtnis. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, **8**, 231-292.
- MASSARO, D. W. (1970a). Perceptual processes and forgetting in memory tasks. *Psychological Review*, **77**, 557-567.
- MASSARO, D. W. (1970b). Retroactive interference in short-term memory for pitch. *Journal of Experimental Psychology*, **83**, 32-39.
- NIPHER, F. E. (1876). Probability of error in writing a series of numbers. *American Journal of Science and Arts* (3rd series), **12**, 79-80.
- NOBLE, C. E. (1952a). An analysis of meaning. *Psychological Review*, **59**, 421-430.
- NOBLE, C. E. (1952b). The role of stimulus meaning (m) in serial verbal learning. *Journal of Experimental Psychology*, **43**, 437-446.
- PHILLIPS, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, **16**, 283-290.
- POSNER, M. I., & KONICK, A. F. (1966). Short-term retention of visual and kinaesthetic information. *Organizational Behaviour and Human Performance*, **1**, 71-85.
- ROSS, H. E., & MURRAY, D. J. (1978). *Translations of E. H. Weber, De subtilitate tactus, 1834, and Der Tastsinn und das Gemeingefühl, 1846*. London: Academic Press.
- SORKIN, R. D. (1966). Temporal interference effects in auditory amplitude discrimination. *Perception & Psychophysics*, **1**, 55-58.
- STIGLER, S. M. (1978). Some forgotten work on memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 1-4.
- SWETS, J. A., TANNER, W. P., & BIRDSALL, T. G. (1961). Decision processes in perception. *Psychological Review*, **68**, 301-340.

- TANNER, W. P. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences*, **89**, 752-765.
- TAYLOR, M. M., & SMITH, S. M. (1975). Monaural detection with contralateral cue. V. Interstimulus interval in MDCC and amplitude discrimination. *Journal of the Acoustical Society of America*, **57**, 1500-1511.
- WARD, L. B. (1937). Reminiscence and rote learning. *Psychological Monographs*, **49**(4, Whole No. 220).
- WEBER, E. H. (1834). *De pulsus, resorptione, auditu et tactu. Annotationes anatomicae et physiologicae*. Leipzig: Koehler. [*The sense of touch* (H. E. Ross & D. J. Murray, Trans.). London: Academic Press, 1978.]
- WEBER, E. H. (1846). Der Tastsinn und das Gemeingefühl. In R. Wagner (Ed.), *Handwörterbuch der Physiologie* (Vol. 3, pp. 481-588). Brunswick: Vieweg. [*The sense of touch*. (H. E. Ross & D. J. Murray, Trans.). London: Academic Press, 1978.]
- WICKELGREN, W. A. (1969). Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, **6**, 13-61.
- WOLFE, H. K. (1886). Untersuchungen über das Tongedächtniss. *Philosophische Studien*, **3**, 534-571.
- WOODWORTH, R. S., & SCHLOSBERG, H. (1955). *Experimental psychology* (2nd ed.). London: Methuen.

(Manuscript received September 25, 1984;
revision accepted for publication February 5, 1985.)