Memory & Cognition
1985, 13 (6), 511-521

An analysis of the strength-latency relationship

BENNET B. MURDOCK, JR.
University of Toronto, Toronto, Ontario, Canada

To account for latency in recognition memory, strength theory assumes that latency decreases symmetrically on both sides of the yes-no criterion. Some of the standard criticisms of the theory are presented. To evaluate these criticisms, explicit expressions for the latency distribution and its mean and variance were obtained. On analysis, the criticisms seem to be unwarranted. However, a satisfactory version of strength theory is not simple, and as many as a dozen or so parameters may be required to account for the data.

Strength theory (Norman & Wickelgren, 1969; Wickelgren & Norman, 1966) has been one of the major theories of recognition memory. Derived from signal-detection theory, it assumes that recognition-memory decisions are based on memory-trace strength. The trace strengths of old and new items are normally distributed, and a yes-no criterion partitions the space. Observations (i.e., trace strengths) falling above the criterion (hits for the old-item distribution; false alarms for the new-item distribution) lead to "yes" responses, and observations falling below the criterion (misses for the old-item distribution; correct rejections for the new-item distribution) lead to "no" responses. For a fuller account, see McNicol and Stewart (1980) or Murdock (1980).

To account for latency data, it is necessary to assume a transfer function that maps strength into latency. That is, for every particular strength value there has to be an associated latency. It generally is assumed (e.g., McNicol & Stewart, 1980; Murdock & Dufty, 1972; Norman & Wickelgren, 1969; Pike, 1973) that latency decreases symmetrically on either side of the yes-no criterion. These relationships are illustrated in Figure 1. The most reasonable transfer function is exponential (Murdock, 1974, p. 282).

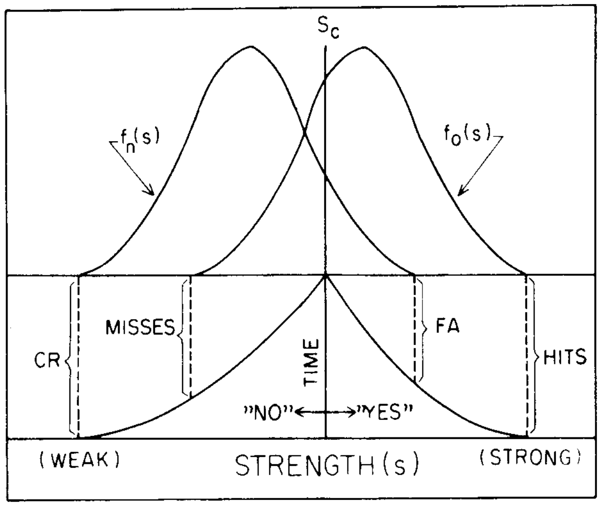

Figure 1. Diagram of the basic strength-latency relationship. The old- and new-item strength distributions are shown at the top, and the transfer function is shown below for correct rejections (CR), misses, false alarms (FA), and hits.

Figure 2 provides a detailed view of the strength-latency transfer function. This figure shows a single strength distribution (in this case, the new-item distribution), the latency transfer function $t = f(s)$, and the resulting latency distribution $f(t)$. Every point on the transfer function is weighted by a strength density to obtain the resulting latency distribution.

There are at least five problems that this model encounters. The first problem is the linearity problem. Reaction-time functions should be a linear function of set size or lag for both hits ("yes" responses to old items) and correct rejections ("no" responses to new items). Below memory span, the linearity occurs when set size is varied over a range of 1-6 items and the probe is a single old or new item (Sternberg, 1966). Above memory span, the linearity occurs for a wide range of lags (old items) or test positions (new items) when a study-test procedure is used (Murdock & Anderson, 1975). Strength theory can clearly predict the increase in reaction time. As set size or lag increases, the average strength of old items will decrease, so more observations will be near the criterion and reaction time will be slower. The question is whether strength theory can predict that the increase in reaction time should be linear.

The second problem is the error variance. The relationships among hits, false alarms, correct rejections, and misses are shown in Figure 1. It would seem that the strength distributions for errors, being more truncated than the distributions for correct responses, should have less variable latency distributions. In a study-test procedure, Murdock and Dufty (1972) found that the latency variance for false alarms was generally greater than for hits

This work was supported by Natural Sciences and Engineering Research Council of Canada Grant No. APA 146. Requests for reprints should be sent to B. Murdock, Department of Psychology, University of Toronto, Toronto, ONT, Canada M5S 1A1. I would like to thank Bill Hockley and Steve Lewandowsky for many helpful comments on the manuscript. Also I would like to thank Richard Shiffrin and his reviewers for their very helpful criticisms and comments on an earlier version.

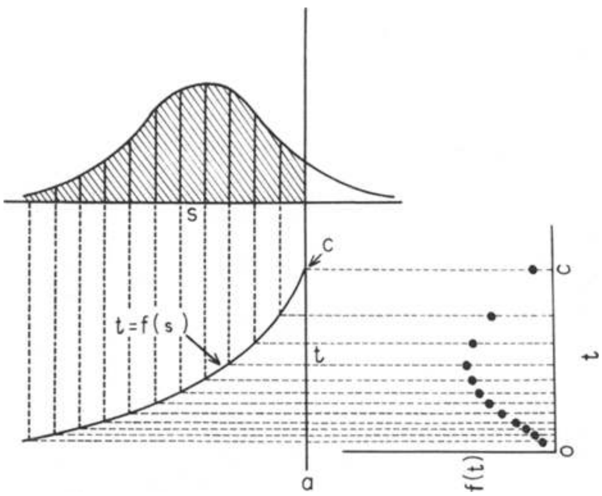

Figure 2. Detailed picture of the strength-latency transfer function. Strength (s) is normally distributed with criterion a. An exponential transfer function maps strength into latency (t), and the resulting latency distribution f(t) is shown by the large dots.

and the latency variance for misses was generally greater than for correct rejections. These results seem to be counter to the theory.

The third problem is the distributional problem. Evidence from a variety of studies (see, e.g., Hockley, 1982; Murdock, 1980) shows that latency distributions are well fit by a convolution of a normal and an exponential distribution (Hohle, 1965; Ratcliff & Murdock, 1976). The latter authors claim that "it is difficult, if not impossible, to get reasonable looking distributions [from strength theory] for both error and correct responses with the same parameter values" (p. 207).

The fourth problem is a variant of the second problem. It may be called the "dominant-no" problem. As noted by Pike (1973), when "no" responses are dominant, false alarms are faster than hits. As Figure 1 should make clear, regardless of criterion location, the model would seem to predict that hits should be faster than false alarms.

The fifth and final problem deals with the relationship between accuracy and latency. As argued in Aubé and Murdock (1974) and in Pike, Dalgleish, and Wright (1977), ability to fit accuracy and latency data simultaneously is a strong test of a model. One of the real strengths of strength theory is that, in principle, it can be applied simultaneously to accuracy and latency data. Unfortunately, there is at least some evidence (Murdock, 1974, p. 283) that parameter values which do in fact quite nicely characterize latency functions for the study-test procedure are well off the mark when applied to accuracy data.

However, none of these problems is as definitive as one might wish. Explicit expressions or extensive computer simulations are necessary if one is to make a convincing analysis. It turns out that explicit expressions for the first two moments (mean and variance) of the latency distribution can be derived. Thus, one can compute predicted mean latency and variability for hits, correct rejections, false alarms, and misses. Also, an expression for the latency distribution itself can be obtained. In the next section, the analysis is presented. In the following section, the problems noted above are reconsidered.

ANALYSIS

The underlying strength distributions are assumed to be normally distributed. The parameters of the old-item distribution are μ_O (the mean of the old-item distribution) and σ_O (the standard deviation of the old-item distribution); the new-item distribution is assumed to have mean zero with standard deviation σ_N. For reasons given above, the strength-latency transfer function for "no" responses is assumed to be a decreasing exponential function of the form $t = ce^{b(s-a)}$, where a is the cutoff on the strength distribution partitioning "yes" and "no" responses, c is the intercept, and b is the rate constant of the transfer function. The transfer function for "yes" responses is $t = ce^{b(a-s)}$ (see Figure 3). The latency distribution f(t) is the distribution whose moments we want; these moments are its mean $\mu_\pm(t)$ and its variance $\sigma^2_\pm(t)$. As Figure 2 shows, f(t) is obtained by multiplying each point on the transfer function by the corresponding strength density.

The mean latency $\mu_\pm(t)$ is:

$$\mu_\pm(t) = c\Phi_{(c)}(z \pm u)\exp(.5u^2 \pm uz)/\Phi_{(c)}(z), \quad (1)$$

where $u = b\sigma$ and $z = (a-u)/\sigma$. The symbol "$\pm$" is to be interpreted as "+" for positive responses (hits and false alarms) and "−" for negative responses (correct rejections and misses). In the expression for u and z, μ and σ refer to the mean and standard deviation of the relevant

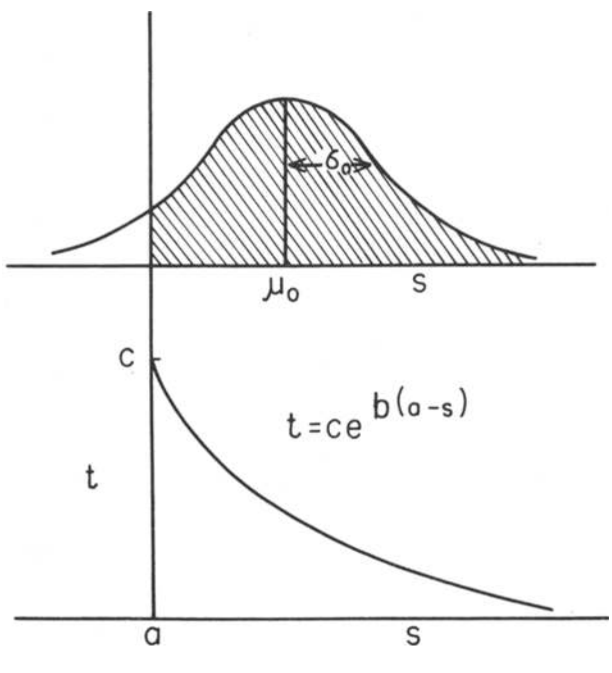

Figure 3. Strength-latency transfer function for old items.

strength distribution (old or new, as the case may be). The symbol $\Phi_{(c)}(z)$ should be read as $\Phi(z)$ for negative responses and as $\Phi_c(z)$, where $\Phi_c(z) = 1-\Phi(z)$, for positive responses. By $\Phi(z)$, I mean the area under the unit normal curve from $-\infty$ to z, and it may be found from tables of the normal curve.

The variance $\sigma^2_{\pm}(t)$ is T1(T2−T3), where

$$T1 = c^2\exp(u^2 \pm 2uz)/\Phi_{(c)}(z),$$

$$T2 = \Phi_{(c)}(z \pm 2u)\exp(u^2), \text{ and}$$

$$T3 = \Phi^2_{(c)}(z \pm u)/\Phi_{(c)}(z), \quad (2)$$

where the same conventions apply. The derivations are presented in the Appendix.

The results for a numerical example are shown in Table 1. For this example, a=1.1, b=1.5, c=1.3, $\mu_N=0.0$, $\sigma_N=0.9$, $\mu_o=1.8$, and $\sigma_o=0.8$. The three 2×2 tables in Table 1 show proportions (proportion of correct rejections and false alarms for new items; proportion of misses and hits for old items), $\mu_{\pm}(t)$, and $\sigma^2_{\pm}(t)$.

The latency distribution $f_{\pm}(t)$ is:

$$f_{\pm}(t) = \exp\{-.5[z \pm u^{-1}(-\ell n(t/c))]^2\}/\sqrt{2\pi}ut\Phi_{(c)}(z), \; o \leq t \leq c. \quad (3)$$

Sample distributions for correct rejections for five values of a (0, 0.5, 1.0, 1.5, and 2.0) and four values of b (0.5, 1.0, 1.5, and 2.0) are shown in Figure 4. In all cases, $\mu_N=0$, $\sigma_N=1.0$, and c=1.0. As can be seen, the shapes of the distributions vary considerably depending upon the values of a and b. The modal value [i.e., that value of t at which f(t) is maximum] is $ce^{-u^2 \pm uz}$ (again, + for positives, − for negatives).

A blowup of three of these distributions is shown in Figure 5. In all three cases, b=1.0. The purpose of Figure 5 is to show that the front end of the distribution drops back to zero for very small values of t. This is always the case, although it is not apparent in Figure 4 for reasons of scale. Thus, although sometimes $f_{\pm}(t)$ is a very good approximation to an exponential or waiting-time distribution, it is not perfect. For the last column in Figure 4 (i.e., b=2.0), the modal values are .018, .007, .002, .0009, and .0003 for a=0.0, 0.5, 1.0, 1.5, and 2.0, respectively.

Table 1
Predicted Proportions, $\mu_{\pm}(t)$, and $\sigma^2_{\pm}(t)$ for a Given Set of Parameter Values

	Proportions		$\mu_{\pm}(t)$		$\sigma^2_{\pm}(t)$	
	Response		Response		Response	
Items	No	Yes	No	Yes	No	Yes
New	.889	.111	.314	.791	.088	1.134
Old	.191	.809	.753	.430	.356	.104

Note—$a=1.1$, $b=1.5$, $c=1.3$, $\mu_N=0$, $\sigma_N=0.9$, $\mu_o=1.8$, and $\sigma_o=0.8$.

DISCUSSION

The five specific problems listed in the introduction are essentially different aspects of a more general problem: whether strength theory can fit all aspects of the data at

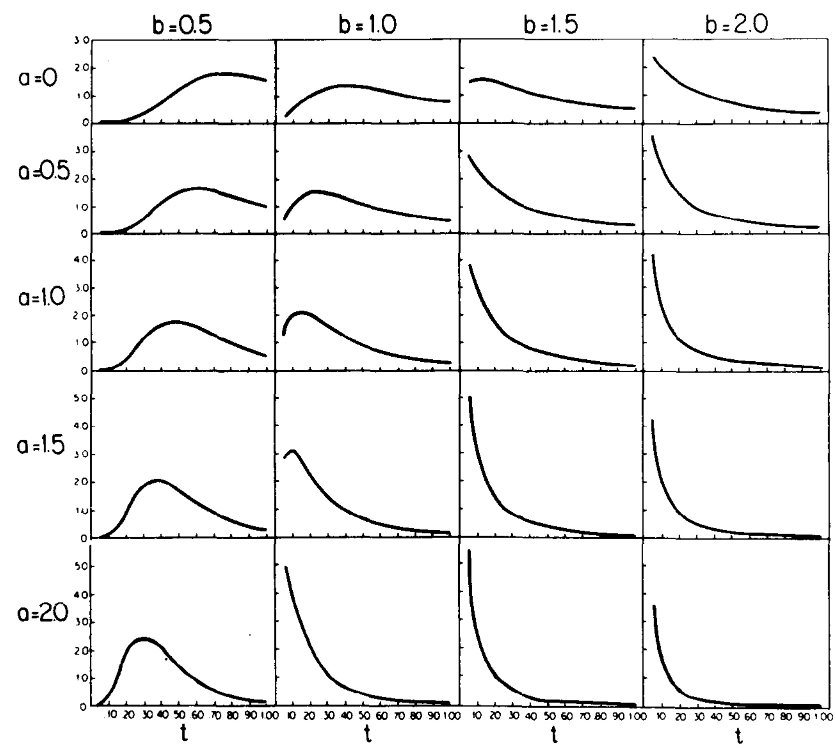

Figure 4. Latency distributions f(t) for five values of a and four values of b.

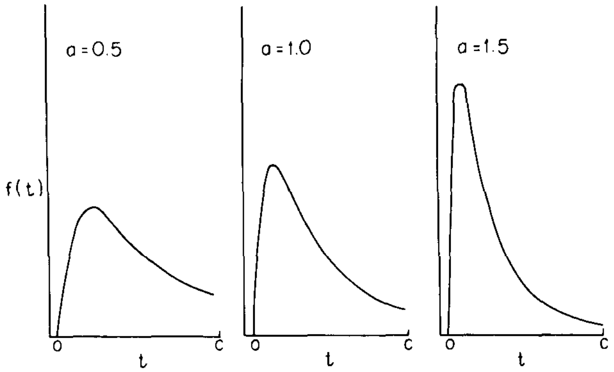

Figure 5. Detailed view of three latency distributions differing in the location of the criterion a. In all three cases, b = 1.0.

a quantitative level. These data would be accuracy and latency data for old and new items for different set sizes or list lengths from different serial positions or test positions. The accuracy data would be proportions of correct and incorrect responses under all possible conditions; the latency data would be means, variances, and distributions for all types of responses under all possible conditions.

At our present stage of knowledge, it seems unlikely that strength theory, or any theory, can do all this perfectly. However, it is worth making the attempt because we will find out just how complex strength theory must be to do a reasonable job, and we will also find out just where and how the theory breaks down. Only in this way can we hope to make any theoretical progress. The data are solid and reliable enough to justify the expenditure of effort, and having the necessary analytic expressions available makes the whole enterprise feasible.

Sternberg Paradigm

Let us start with the Sternberg paradigm, since it is simpler than the study-test paradigm. It is quite clear that a simple strength model will not work. By a simple strength model, I mean a model in which there is one new-item distribution and p (for set size) old-item distributions, a single criterion, and the same strength-latency transfer function for all set sizes. Following strength theory (Wickelgren & Norman, 1966), there would be exponential decay of d′; the mean strength of the last item would be greatest, of the next-to-last item next greatest, and so on, with an exponential decline. Also, equal variances would be assumed.

Why will this simple strength theory not work? If the criterion stayed constant as set size varied, then the mean reaction time for negatives would not vary with set size, and this is clearly false. Suppose the criterion decreased as set size increased. This is not impossible; in either a fixed-set or a varied-set procedure, subjects know the set size before the probe is presented, so the location of the criteria could be adjusted accordingly. This adjustment could produce an increase in reaction time for negatives, but then there would be a problem for positives. The mean reaction time for the last item in the list would decrease as set size increased. In general, it does not; it either stays the same or increases slightly.

However, there are more complicated versions of strength theory which do a much better job. In fact, it is quite possible to have a version of strength theory that generates linear and parallel functions for old and new items with reasonable error rates. An example is shown in Table 2. Predicted reaction time for new items is shown in the last column, and predicted mean reaction time for old items is shown in the next-to-last column. The data I fitted were ideal data; slopes of 35 msec/item for both old and new items, an intercept for old items of 700 msec, and an intercept for new items of 750 msec. Although not shown, the mean accuracy at each set size was at least 95% and the highest error rate was 12% at Serial Position 1 of Set Size 6. Essentially, this is "error-free" performance when averaged over serial position.

To achieve these results, it was necessary to let five parameters vary across conditions. These parameters were μ_O and σ_O, the mean and standard deviation of the old-item distribution; σ_N, the standard deviation of the new-item distribution; b, the rate constant of the exponential transfer function; and the intercept c, whose value was higher for negatives than for positives, although this did not vary with set size.

For those who are interested, here are the numerical values. The mean of the old-item distribution was 6.15 for the last serial position in each set size, and decreased exponentially with rate constant .084 with each earlier item. The old-item variance was always a constant proportion of the mean, and this proportion was .440. The new-item variance increased with set size as the sum of the geometric series 1, α^2, α^4, . . . , α^{2p} for Set Size p. The parameter α is the same α as in Equation 1 of Murdock (1982, 1983) and is the serial position constant c_j in Anderson (1973); the value of α here was .706. The slope of the transfer function b decreased exponentially with set size. The value of b at Set Size 1 was .706, and the rate constant for the exponential decay was .314. The location of the criterion a was the same for all set sizes; the value of a was 2.44. The value of c, the intercept of the strength-latency transfer function, was 306 msec for positives and 367 msec for negatives. The predicted values from Equation 1 for the parameter values were augmented

Table 2
Predicted Latency for Old and New Items by Set Size and Serial Position

Set Size	Serial Position 1	2	3	4	5	6	Old Items Mean	New Items
1	738						738 (735)	781 (785)
2	774	760					767 (770)	819 (820)
3	818	803	788				803 (805)	856 (855)
4	863	849	834	819			841 (840)	892 (890)
5	903	891	878	865	851		878 (875)	926 (925)
6	935	926	916	905	893	881	909 (910)	956 (960)

Note—Fitted data in parentheses.

by 700 msec as a "time for other stages" (TOS), so this fit characterized the comparison process only. Overall, the fit was excellent; the standard error (root-mean-square value, the square root of the mean sum of squared deviations between predicted and observed data points) was only 2.4 msec.

This particular version of strength theory is by no means the only version which will give a good fit. I have found a number of others that are almost as good, if not as good, in fitting the set size functions. The problem is the serial-position effects. This model does not do a very good job fitting serial-position data, but the other versions are worse. The predicted serial-position curves for this version of strength theory are also shown in Table 2.

In case it is not clear, the parameter estimation program (SIMPLX; Nelder & Mead, 1965) optimized the parameters for the ideal set-size functions. The old- and new-item values are shown in parentheses in Table 2. These parameters were used to generate the predicted serial-position effects for old items; these are the serial-position data shown in Table 2.

Basically, there seem to be two problems. First, the serial-position curves are are not bowed enough. Second, the increase in the predicted latency for the last serial position as set size increases is too great. These problems may be due to the fact that we fit the set-size functions directly and then used the obtained parameter values to predict the serial-position functions. The remedy would seem to be to fit the serial-position curves directly.

A brief comment is in order before describing this attempt. In modeling (and thinking about) data from the Sternberg paradigm, one generally tries to explain the linear set-size functions and then cope with the serial-position effects as best one can. An alternative approach may be suggested. Perhaps one should try to explain the serial-position effects directly, then derive the linear set-size functions from the serial-position effects. This approach was suggested by Murdock (1971), and is explicit or implicit in many other accounts of Sternberg data. We shall apply the same reasoning to strength theory. Can it explain the serial-position effects? If so, then the set-size functions must necessarily follow (Franklin, 1980).

As a test, we took the data from Murdock and Franklin (1984), which shows very substantial serial-position effects. Because there were primacy effects as well as recency effects in these data, we decided to let μ_o vary with both distance from the beginning and distance from the end of the list. That is, the first component of μ_o decreased with the number of prior items and, thus, represented a proactive inhibition (PI) effect. The second component of μ_o decreased with the number of subsequent items and, thus, represented a retroactive inhibition (RI) effect. Specifically,

$$\mu_o(p,j) = a_1\exp(-b_1 j) + a_2\exp[-b_2(p-j)],\ j = 1,2,\ldots,p, \tag{4}$$

where a_1 and b_1 are the intercept and rate constant of the PI component, a_2 and b_2 are the intercept and rate constant of the RI component, and j denotes serial position in Set Size p. Although to some extent, Equation 4 is arbitrary, it essentially formalizes PI and RI effects which, must be represented in any serious model.

In some preliminary fits, b_1 was consistently quite large, so to save a parameter we arbitrarily set it to seven times a_1. This results in a very abbreviated (essentially one-item) primacy effect. The other free parameters were the old-item variance ratio (the ratio of the variance to the mean), the location of the criterion a, the starting value and rate constant for b, the slope and rate constant of the strength-latency transfer function, its intercept c, and TOS. The intercept c was the same for positive and negative responses, but a constant "extra time for negatives" (ETFN) was added to the predicted latency to account for the intercept difference. ETFN was a fixed parameter and was the same for all set sizes.

In computing the standard error, we weighted each set size equally and, within each set size, positives and negatives equally. This followed the experimental design, in which each set size was tested equally often. Within each set size, the probability of an old- or new-item probe was always .5, and the serial position of old-item probes was randomly selected.

Also, we fit both accuracy and latency data simultaneously. To compute an overall goodness-of-fit measure, we had to establish some correspondence between accuracy and latency. Initially we used an accuracy-to-latency (ATL) scale value of 1.0 where 1 error in 1,000 was equal to 1 msec. In the final fit, this was changed to .1, as subjects in the experiment were required to maintain a high level of accuracy. It is probable that errors often resulted from extraneous factors. Admittedly, any ATL is arbitrary, but some value has to be used if one is to fit accuracy and latency simultaneously.

Table 3
Predicted Latency and Accuracy for Old and New Items by Set Size and Serial Position

	Serial Position						
Set Size	1	2	3	4	5	6	New Items
			Latency				
2	455	462					588
	(462)	(480)					(554)
4	512	540	510	485			629
	(521)	(552)	(526)	(477)			(658)
6	631	690	643	598	558	523	691
	(611)	(697)	(641)	(609)	(521)	(496)	(687)
			Accuracy				
2	.011	.016					.022
	(.028)	(.026)					(.008)
4	.031	.048	.029	.016			.022
	(.090)	(.060)	(.008)	(.024)			(.050)
6	.069	.116	.077	.049	.029	.016	.022
	(.080)	(.040)	(.006)	(.013)	(.013)	(.000)	(.050)

Note—Obtained results in parentheses. Data from Murdock and Franklin (1984).

For the best-fitting parameter values, the latency component of the standard error was 21 and the accuracy component was 29. Thus, the fits are good to within about 20 msec and 3% error. Observed and predicted results (latency and accuracy) are shown in Table 3. The best-fitting parameter values were $a_1=.753$, $a_2=4.971$, $b_2=.097$; the old-item variance ratio was .384; the criterion a was 2.013; the starting value of b was 2.44; the rate constant was .166; the intercept c was 941 msec; and TOS was 433 msec. As a fixed parameter, ETFN was set to 94 msec, the intercept difference in the set-size functions for positives and negatives.

To summarize, the best fit showed that RI is both much larger and more persistent than PI, although both are required to explain the bow-shaped serial-position curve. The old-item variance changed with serial position, but the new-item variance was set to 1 for all set sizes. The criterion seems to be constant across set size. (It did not help the fit to let the criterion vary with serial position.) Finally, the slope of the strength-latency transfer function decreased exponentially with set size.

What does it mean to say that the slope of the transfer function decreased with set size? Unfortunately, no ready answer is forthcoming. All I can say is that the slope of this function is one of the parameters of the model, and a decrease in its value affects both the mean and variance of the latency distribution. What it means from the point of view of a psychological process is not clear, as strength theory is an analytic model, not a process model.

One of the problems with strength theory is the dominant-no problem. If "no" responses are dominant, how can the latency for hits be greater than the latency for false alarms? There is at least one answer; namely, σ_o must be less than σ_N. As shown in Table 4, for $a=1.8$ and for $a=2.0$, when $\sigma_o=.25$, hits are slower than false alarms. Other parameter values here were $\mu_o=2.0$, $b=1.0$, and $c=1.0$. Of course, strength theory must explain why the variances are the way they are, but that is another matter.

Study-Test Paradigm

How well does strength theory work when applied to the study-test procedure? In this procedure, accuracy and latency covary, and errors range from 0 to 25% depending upon study position, test position, and lag. The very fact that accuracy, as well as latency, must be included makes this a much more challenging problem than the Sternberg paradigm.

Table 4
Mean Latency for Hits as a Function of μ_o and Hits and False Alarms as a Function of a

	Hits	
σ_o	$a=1.8$	$a=2.0$
3	.267	.284
2	.319	.333
1	.494	.524
0.5	.652	.697
0.25	.760	.828
0.125	.812	.908
False Alarms	.777	.853

Note—$\mu_o=2.0$, $b=c=1.0$.

One needs two criteria if one is to accommodate confidence-judgment data: a, a lower criterion for high-confident correct rejections and misses, and b, an upper criterion for high-confident hits and false alarms (see Figure 1 in Murdock & Anderson, 1975). With the new-item variance fixed at 1.0, it was necessary to let the lower criterion a decrease over trials to accommodate the decrease in the proportion of high-confident correct rejections. For serial-position effects, I still used Equation 4, but the RI component let study and test interference vary separately. There is conflicting evidence here (Norman & Waugh, 1968), and I was not sure what to do. For study-test data, I selected the data from Table 3 of Murdock and Anderson (1975), as this was representative data from a few practiced subjects given intensive testing.

Altogether there were 14 parameters. Nine [a_1, b_1, a_2, b_{2s} (study), b_{2t} (test), the old-item variance ratio, a (starting value), a (rate constant), and b] were sufficient to determine the accuracy. Five (the starting value and rate constant of b, the slope of the strength-latency transfer function, its intercept c, TOS, and ETFN) affected latency but did not affect accuracy. Because our version of SIMPLX allows only 9 free parameters, I did some exploratory estimations and, as a result, fixed 5 of these parameters in advance and did the final estimation on only 9 free parameters. These were a_1, b_1, b_{2s}, the old-item variance ratio, the starting value and rate constant of a, the lower criterion, the starting value of b, the slope of the transfer function, its intercept c, and TOS. The parameters I fixed in advance were $a_2=8.3-a_1$, $b_{2t}=b_{2s}/4$, the upper criterion $b=4.9$, and the rate of change of the slope of the transfer function equal to 1/16th of its starting value. Admittedly these are arbitrary, but some way had to be found to reduce 14 free parameters to 9 free parameters.

The best-fitting parameter values were 5.245, .059, .353, .178, 1.239, .041, 2.706, 284, and 635 for a_1, b_1, b_{2s}, the old-item variance ratio, the starting value and rate constant for the lower criterion a, the starting value for the slope b of the transfer function, its intercept c, and TOS, respectively. The fit was really quite good, especially considering that 120 data points were fit by nine free plus five fixed parameters. The standard error was 17 msec for latency and 1.9% for accuracy. Observed and predicted latencies and proportions are shown in Table 5, broken down by study and test position (blocks of three) for old items and by test position for new items.

With one exception, these parameter values were comparable to those from the Sternberg paradigm. The exception was b_1. In the Sternberg paradigm, it was very large; in the study-test paradigm, it was almost zero. The implication is that, in the Sternberg paradigm, the primacy effect is due to some special status of the first item. Even the second item does not share this benefit. In the study-

Table 5
Predicted and Observed Latencies and Proportions for High-Confident Hits and Correct Rejections by Input and Output Blocks

Input Block	Output Block 1	2	3	4	5	6	7	8	9	10
				Observed Latency						
1	723	671	664	675	727	728	721	769	787	866
2	733	668	693	704	748	730	756	792	798	817
3	714	685	683	673	704	754	766	764	807	790
4	668	646	657	693	703	739	763	745	768	800
5	637	634	657	674	678	715	730	768	785	797
New Items	931	720	729	736	773	787	796	807	809	840
				Predicted Latency						
1	670	680	692	705	719	734	750	766	781	796
2	670	681	693	707	723	739	756	772	789	804
3	665	676	689	704	721	739	757	775	792	808
4	656	666	679	694	712	731	751	771	789	807
5	645	652	663	677	694	715	736	758	779	799
New Items	716	725	735	747	760	775	791	807	823	839
				Observed Proportion						
1	.848	.840	.848	.877	.807	.826	.803	.729	.780	.739
2	.857	.856	.828	.821	.792	.747	.753	.693	.700	.672
3	.880	.874	.828	.839	.790	.775	.687	.721	.671	.718
4	.931	.904	.896	.876	.820	.812	.794	.731	.693	.698
5	.976	.981	.951	.931	.900	.889	.815	.780	.820	.759
New Items	.908	.878	.870	.847	.833	.838	.830	.852	.811	.817
				Predicted Proportion						
1	.854	.841	.828	.816	.804	.793	.782	.772	.762	.753
2	.857	.838	.820	.801	.783	.766	.749	.733	.718	.703
3	.887	.864	.841	.816	.792	.767	.743	.720	.697	.675
4	.936	.915	.891	.864	.836	.806	.775	.744	.713	.683
5	.979	.967	.951	.930	.905	.877	.844	.810	.773	.736
New Items	.892	.883	.873	.863	.853	.843	.833	.824	.814	.804

test procedure, by contrast, there is a special status afforded to every study item, and it does not diminish with number of prior items. Strength of an item, then, has two components. The first component results only from list presentation, and persists unchanged throughout the study and test phase. The second component, slightly smaller in magnitude, falls off the more items that follow. The RI effect is greater for study than for test items.

It would seem, then, that strength theory can fit accuracy and latency data with the same set of parameters and, moreover, can do so very well. The standard errors are really quite small. Admittedly, this is a more complex version of strength theory than is usually considered, but rather than abandon a theory that has weaknesses, it is sometimes better to try to fix it up. Perhaps the proper conclusion to the original problem is to realize that a simple version of strength theory indeed cannot handle accuracy and latency with the same set of parameter values, but a more complex version of the theory can.

Another of the problems of strength theory mentioned in the introduction was the error variance. It was claimed in Murdock and Dufty (1972) that, according to strength theory, errors should be less variable than correct responses. This claim is wrong. There are parameter values such that errors should be less variable, but there are other parameter values such that errors should be more variable. To illustrate, Table 6 presents the results of a four-parameter grid search. There were three levels of a (1.0, 1.5, and 2.0), three levels of b (0.5, 1.0, and 1.5), three levels of σ_o (0.5, 1.0, and 1.5), and four levels of μ_o (1.0, 1.5, 2.0, and 2.5). In all cases, $\mu_N = 0$ and $\sigma_N = 1.0$. Table 6 shows the variance of correct rejections and the variance of misses over this four-parameter space.

To summarize these results, the number of cases where the variance of correct rejections was less than the variance of misses increased with a, b, and σ_o, but decreased with μ_o. The effect of σ_o seems particularly clear. For $\sigma_o = 1.5$, in 35 out of 36 cases the variance of correct rejections was less than the variance of misses. Thus, one of the conclusions of Murdock and Dufty (1972) must be retracted. The fact that, in the data, errors were more variable than correct responses does not constitute negative evidence for strength theory.

Another problem with strength theory is the distribution analysis. Empirically, many reaction-time distributions are well represented by the convolution of a normal and an exponential distribution (Ratcliff & Murdock, 1976). Can strength theory accommodate this result? Given the parameter values we find, it seems likely that the decision state must represent the exponential component (see Figure 4). Thus, if the other stages are normally distributed, then the distribution analysis is not a problem for strength theory.

It has been mentioned that the distribution f(t) always drops back to zero for low values of t. Thus, f(t) could not be a pure exponential distribution. However, this imperfection could be obscured by criterion variability. If the location of a, the criterion, varied somewhat from trial to trial, then the leading edge of f(t) would be blurred. This possibility was investigated by Monte Carlo methods, and sample distributions are shown in Figure 6. Here the criterion was centered on a and its variability (standard deviation) is η. In the two left-hand panels $\eta=0$, so there was no criterion variability; however, in the two right-hand panels $\eta=1$, so the criterion variability was the same as that of the new-item distribution.

This Monte Carlo simulation was based on 1,000 trials, and the bin width for these frequency polygons was 25 msec. These distributions (polygons) were for correct rejections, and both b and c were set to 1. The values of both a and η affect the resulting distribution, but the important point to note is that criterion variability seems to obscure the initial rise in f(t) at very low values of t. Thus, even some of the dubious exponentials in Figure 4 might be better if criterion variability existed.

CONCLUSIONS

The problems noted in the introduction turn out not to be problems at all. Strength theory can fit the data, and it does an excellent job. It can fit accuracy and latency simultaneously with the same set of parameter values; it can generate linear set-size functions and very respectable serial-position and study-test functions. Error variability is no problem, and the latency distributions seem to be generally exponential. Although the fits to the data are not perfect, they are limited by noise in the data and it seems unlikely any other theory can do appreciably better.

The advantages of strength theory should not be minimized. It allows one to account quantitatively for accuracy and latency, albeit with a fair number of parameters. Explicit expressions for latency can be derived, and the derivations are not particularly difficult. One can argue that the number of free parameters is excessive, but consider what predictions are possible. One can predict accuracy proportions, latency means, variance, and distributions for correct rejections, misses, hits, and false alarms for all serial positions in the Sternberg paradigm and all study-test positions in the study-test paradigm, all with the same set of parameter values. Standard methods of parameter estimation, such as SIMPLX or STEPIT, can be used to fit the model to data or, given particular parameter values, specific predictions can be made.

The fact that strength theory predicts a number of aspects of the data simultaneously means that everything is interrelated. One soon comes to appreciate what this means when one tries to fit the model to a given set of data. If one changes the location of the criterion to increase accuracy, one also changes the latency. If one changes the mean or variance of the old-item distribution to adjust, for example, hits, then one also changes misses. If the slope of the strength-latency transfer function changes, this will affect both means and variances of posi-

Table 6
Variance of Correct Rejections (Column 4) and Misses (Columns 5–8) as a Function of a, b, σ_o, and μ_o

				Misses			
a	b	σ_o	CR	$\mu_o=1.0$	$\mu_o=1.5$	$\mu_o=2.0$	$\mu_o=2.5$
1.000	0.500	0.500	0.041	0.014	0.008	0.005	0.003
1.000	0.500	1.000	0.041	0.034	0.029	0.024	0.019
1.000	0.500	1.500	0.041	0.051	0.047	0.043	0.039
1.000	1.000	0.500	0.059	0.034	0.024	0.016	0.011
1.000	1.000	1.000	0.059	0.063	0.058	0.052	0.045
1.000	1.000	1.500	0.059	0.074	0.074	0.072	0.069
1.000	1.500	0.500	0.058	0.051	0.039	0.028	0.026
1.000	1.500	1.000	0.058	0.074	0.073	0.069	0.067
1.000	1.500	1.500	0.058	0.077	0.083	0.081	0.234
1.500	0.500	0.500	0.040	0.019	0.014	0.008	0.005
1.500	0.500	1.000	0.040	0.039	0.034	0.029	0.024
1.500	0.500	1.500	0.040	0.054	0.051	0.047	0.043
1.500	1.000	0.500	0.047	0.041	0.034	0.024	0.016
1.500	1.000	1.000	0.047	0.064	0.063	0.058	0.052
1.500	1.000	1.500	0.047	0.071	0.074	0.074	0.072
1.500	1.500	0.500	0.041	0.054	0.051	0.039	0.028
1.500	1.500	1.000	0.041	0.069	0.074	0.073	0.069
1.500	1.500	1.500	0.041	0.069	0.077	0.083	0.081
2.000	0.500	0.500	0.034	0.020	0.019	0.014	0.008
2.000	0.500	1.000	0.034	0.041	0.039	0.034	0.029
2.000	0.500	1.500	0.034	0.055	0.054	0.051	0.047
2.000	1.000	0.500	0.032	0.034	0.041	0.034	0.024
2.000	1.000	1.000	0.032	0.059	0.064	0.063	0.058
2.000	1.000	1.500	0.032	0.066	0.071	0.074	0.074
2.000	1.500	0.500	0.024	0.035	0.054	0.051	0.039
2.000	1.500	1.000	0.024	0.058	0.069	0.074	0.073
2.000	1.500	1.500	0.024	0.060	0.069	0.077	0.083

Note—In all cases, $\mu_N=0$ and $\sigma_N=1.0$.

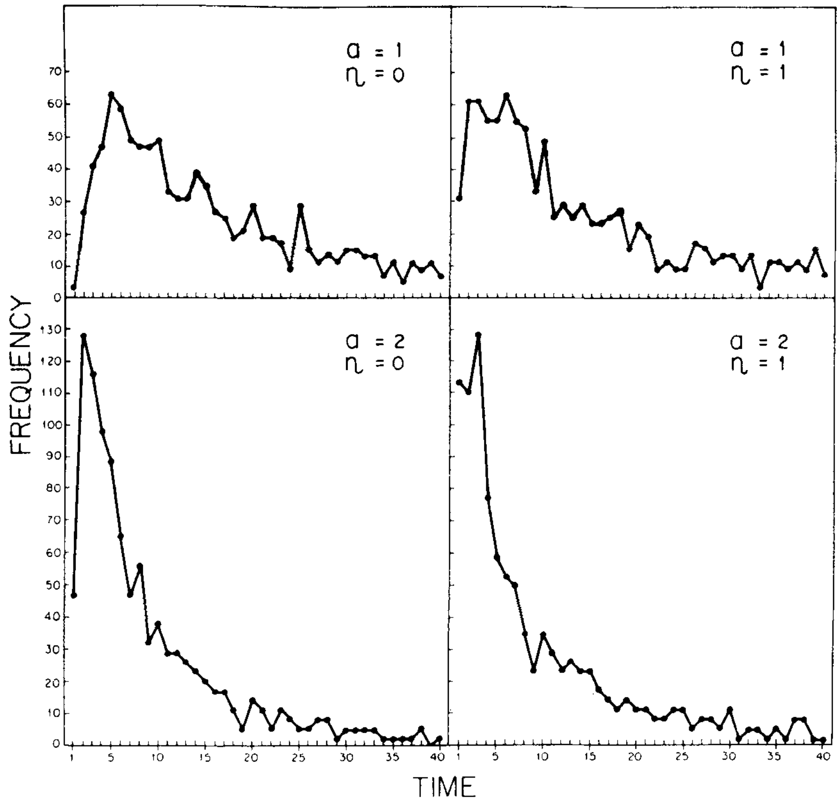

Figure 6. Results of four Monte Carlo computer simulations to show the results of criterion variability. The standard deviation of the criterion distribution is η, and the criterion is centered on a.

tives or negatives. And so it goes; one cannot make local changes to patch one problem without at the same time changing other aspects of the data and perhaps causing new problems.

Although frustrating, this feature of strength theory is probably realistic. As in most complex systems, everything probably is interrelated. The human memory system is a complex system, and any one change may have widespread ramifications. Our theories should face up to these complexities, rather than pretend they do not exist.

On the other hand, the difficulties with strength theory should also be noted. A simple version cannot account for the data, and a version which can account for the data requires some ad hoc assumptions. It requires a dozen or so parameters and, with so many parameters, separation of parameters may not be easy. The parameter space may be quite flat, so fairly substantial changes in the data may not be easily attributable to one or two parameters. Other types of models (e.g., Shiffrin's SAM model; see Gillund & Shiffrin, 1984) may also have this problem, so it is not necessarily unique to strength theory.

Strength theory uses a functionalistic approach, not a structuralistic approach (Pieters, 1983). Admittedly, this is a matter of theoretical preference, but it does seem to be the case that many cognitive psychologists today prefer process models. Strength theory describes a simple transfer relationship, and this makes possible the latency derivations. However, strength theory does not provide many insights as to how or why the decision system operates, or, for that matter, what the origin of the original strength distributions might be. The slope of the transfer function seems to decrease as set size or test position increases. What does this transfer function represent, and why does its slope decrease? Those who want a process view will have to look elsewhere for the answers to these questions.

To summarize, a simple version of strength theory with parameters μ_N, σ_N, μ_O, σ_O, a, b, and c (Figure 3) cannot adequately account for accuracy and latency data from the Sternberg or the study-test procedure. If several of these parameters are allowed to vary across conditions (set size, output, or serial position), then it can. Explicit expressions for the latency distribution, mean, and variance for hits, false alarms, misses, and correct rejections may be obtained by standard methods, and these expressions may be useful in further developments of the theory.

REFERENCES

Anderson, J. A. (1973). A theory for the recognition of items from short memorized lists. *Psychological Review*, **80**, 417-438.

Aubé, M., & Murdock, B. B., Jr. (1974). Sensory stores and high-speed scanning. *Memory & Cognition*, **2**, 27-33.

Franklin, P. E. (1980). *The effect of experimental procedures on memory scanning*. Unpublished doctoral dissertation, York University, Toronto.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.

HOCKLEY, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 497-512.
HOEL, P. G. (1962). *Introduction to mathematical statistics* (3rd ed.). New York: Wiley.
HOHLE, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology*, **69**, 382-386.
MCNICOL, D., & STEWART, G. W. (1980). Reaction time and the study of memory. In A. T. Welford (Ed.), *Reaction times* (pp. 253-307). London: Academic Press.
MURDOCK, B. B., JR. (1971). A parallel-processing model for scanning. *Perception & Psychophysics*, **10**, 289-291.
MURDOCK, B. B., JR. (1974). *Human memory: Theory and data.* Potomac, MD: Erlbaum.
MURDOCK, B. B., JR. (1980). Short-term recognition memory. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 497-519). Hillsdale, NJ: Erlbaum.
MURDOCK, B. B., JR. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89**, 609-626.
MURDOCK, B. B., JR. (1983). A distributed memory model for serial-order information. *Psychological Review*, **90**, 316-338.
MURDOCK, B. B., JR., & ANDERSON, R. E. (1975). Encoding, storage, and retrieval of item information. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Hillsdale, NJ: Erlbaum.
MURDOCK, B. B., JR., & DUFTY, P. O. (1972). Strength theory and recognition memory. *Journal of Experimental Psychology*, **94**, 284-290.
MURDOCK, B. B., JR., & FRANKLIN, P. E. (1984). Associative and serial-order information: Different modes of operation? *Memory & Cognition*, **12**, 243-249.
NELDER, J. A., & MEAD, R. (1965). A simplex method for function minimization. *Computer Journal*, **7**, 308-313.
NORMAN, D. A., & WAUGH, N. C. (1968). Stimulus and response interference in recognition-memory experiments. *Journal of Experimental Psychology*, **78**, 551-559.
NORMAN, D. A., & WICKELGREN, W. A. (1969). Strength theory of decision rules and latency in short-term memory. *Journal of Mathematical Psychology*, **6**, 192-208.
PIETERS, J. P. M. (1983). Sternberg's additive factor method and underlying psychological processes: Some theoretical considerations. *Psychological Bulletin*, **91**, 411-423.
PIKE, R. (1973). Response latency models for signal detection. *Psychological Review*, **80**, 53-68.
PIKE, R., DALGLEISH, L., & WRIGHT, J. (1977). A multiple-observations model for response latency and latencies of correct and incorrect responses in recognition memory. *Memory & Cognition*, **5**, 580-589.
RATCLIFF, R., & MURDOCK, B. B., JR. (1976). Retrieval processes in recognition memory. *Psychological Review*, **83**, 190-214.
STERNBERG, S. (1966). High-speed scanning in human memory. *Science*, **153**, 652-654.
WICKELGREN, W. A., & NORMAN, D. A. (1966). Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, **3**, 316-347.

APPENDIX

For negative responses

$$\mu_{-}t = \int_0^c t f(t)\,dt =$$

$$\frac{1}{\Phi[(a-\mu)/\sigma]}\int_{-\infty}^{a} c e^{b(s-a)}\Phi(s)\,ds =$$

$$\frac{ce^{-ba}}{\Phi[(a-\mu)/\sigma]}\,\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{a} e^{bs-1/2[(s-\mu)/\sigma]^2}ds, \quad \text{(A1)}$$

where $\Phi(s)$ is the new-item strength distribution and $\Phi\,(a-\mu)/\sigma$ is the area under the curve from $-\infty$ to a normalized in terms of μ and σ (see Figure 1).

Equation A1 can be rewritten as

$$\mu_{-}(t) = \frac{k}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{a} e^{-1/2[(s^2-2\mu s-2b\sigma^2 s)/\sigma^2]}ds, \quad \text{(A2)}$$

where

$$k = \frac{ce^{-ba-(\mu^2/2\sigma^2)}}{\Phi[(a-\mu)/\sigma]}. \quad \text{(A3)}$$

If we take the exponent of the integral in Equation A2 and complete the square, we have

$$\mu_{-}(t) = \frac{k'}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{a} e^{-1/2[(s-\mu-b\sigma^2)/\sigma]^2}ds, \quad \text{(A4)}$$

where

$$k' = ke^{1/2[(\mu+b\sigma^2)/\sigma]^2}. \quad \text{(A5)}$$

The expression in Equation A4 is simply $k'\Phi[(a-\mu-b\sigma^2)/\sigma]$, so by combining Equations 3-5 we have

$$\mu_{-}(t) = ce^{b(\mu-a)+(b^2\sigma^2)/2}\,\frac{\Phi[(a-\mu-b\sigma^2)/\sigma]}{\Phi[(a-\mu)/\sigma]}. \quad \text{(A6)}$$

Thus, we have an explicit expression for the mean of the latency distribution for negative responses in terms of the five parameters of the model: μ and σ, the mean and standard deviation of the strength distribution; a, the criterion or cutoff on the strength distribtuion; and b and c, the rate constant and the intercept of the exponential transfer function. Equation A6 would apply to both correct rejections (μ=o) and missess (μ > o). Figure 3 shows the strength-latency transfer function for hits, and the analysis for hits and false alarms is similar.

Since $\mathrm{Var}(t) = \mu_2(t)-\mu^2(t)$, where $\mu_2(t)$ is the second moment, we have

$$\mu_2(t) = \int_0^c t^2 f(t)\,dt =$$

$$\frac{1}{\Phi[(a-\mu)/\sigma]}\int_{-\infty}^{a} c^2 e^{2b(s-a)}\Phi(s)\,ds =$$

$$\frac{c^2e^{-2ba}}{\Phi[(a-\mu)/\sigma]}\,\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{a} e^{2bs-1/2[(s-\mu/\sigma]^2}ds. \quad \text{(A7)}$$

Equation A7 can be rewritten as

$$\mu_2(t) = \frac{h}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{a} e^{-1/2[(s^2-2\mu s-4b\sigma^2 s)/\sigma^2]}ds, \quad \text{(A8)}$$

where

$$h = \frac{c^2e^{-2ba-(\mu^2/2\sigma^2)}}{\Phi[(a-\mu)/\sigma]}. \quad \text{(A9)}$$

Again we complete the square in Equation A8, and we have

$$\mu_2(t) = \frac{h'}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{a} e^{-1/2[(s-\mu-2b\sigma^2)/\sigma]^2}\,ds, \quad (A10)$$

where

$$h' = he^{1/2[(\mu+2b\sigma^2)/\sigma]^2}. \quad (A11)$$

The expression in Equation A10 is simply $h'\Phi[(a-\mu-2b\sigma^2/\sigma]$, so by combining Equations A9-A11, we have

$$\mu_2(t) = c^2e^{2b(\mu-a)+2b^2\sigma^2}\frac{\Phi[(a-\mu-2b\sigma^2)/\sigma]}{\Phi[(a-\mu)/\sigma]}, \quad (A12)$$

from which we can calculate the variance.

It simplifies the expressions and clarifies the relationships involved if we make the substitutions $z=(a-\mu)/\sigma$ and $u=b\sigma$. The we can write

$$\mu_-(t) = c\,e^{(u^2/2)-uz}\frac{\Phi(z-u)}{\Phi(z)}, \quad (A13)$$

and

$$\mathrm{Var}_-(t) = \frac{c^2e^{u^2-2uz}}{\Phi(z)}\left[e^{u^2}\Phi(z-2u) - \frac{\Phi^2(z-u)}{\Phi(z)}\right]. \quad (A14)$$

In a similar way, one can obtain the mean and variance for hits and false alarms. The transfer function is $t=ce^{b(a-s)}$(see Figure 3) and one must replace $\Phi[(a-\mu)/\sigma]$ by $\Phi_c[(a-\mu)/\sigma]$ where $\Phi_c[(a-\mu)/\sigma] = 1-\Phi[(a-\mu)/\sigma]$. Otherwise the development is the same. Then for positive responses (hits and false alarms), we have

$$\mu_+(t) = ce^{(u^2/2)+uz}\frac{\Phi_c(z+u)}{\Phi_c(z)}, \quad (A15)$$

and

$$\mathrm{Var}_+(t) = \frac{c^2e^{u^2+2uz}}{\Phi_c(z)}\left[e^{u^2}\Phi_c(z+2u) - \frac{\Phi_c^2(z+u)}{\Phi_c(z)}\right]. \quad (A16)$$

Thus, this analysis leads to explicit expressions for the first two moments of the latency distribution for all four cells of the 2×2 signal-detection/strength-theory model of recognition memory: correct rejections, misses, hits, and false alarms.

To obtain an expression for the latency distribution itself, one can use the change-of-variable technique (e.g., Hoel, 1962). For negative responses, we must solve the transfer function $t=ce^{b(s-a)}$ for s in terms of t, then $|ds/dt|=1/bt$ and

$$f_-(t) = \frac{1}{\sqrt{2\pi}\,ut}\frac{e^{-1/2[z+(1/u)\ln(t/c)]^2}}{\Phi(z)}, \quad (A17)$$

where z and u are as defined above. For positive responses, the transfer function is $t=ce^{b(a-s)}$ and

$$f_+(t) = \frac{1}{\sqrt{2\pi}\,ut}\frac{e^{-1/2\{z-(1/u)[\ln(t/c)]\}^2}}{\Phi_c(z)}, \quad (A18)$$

where as above $\Phi_c(z)=1-\Phi(z)$.

(Manuscript received June 18, 1984;
revision accepted for publication November 15, 1985.)