Journal of Experimental Psychology:
Human Perception and Performance
1992, Vol. 18, No. 2, 530-549

0096-1523/92/$3.00

On the Time Course of Perceptual Information That Results From a Brief Visual Presentation

Geoffrey R. Loftus, Janine Duncan, and Paul Gehrig
University of Washington

A briefly presented visual stimulus engenders an *available-information function* that lags behind the physical stimulus. We report two experiments that focus on the *iconic-decay* portion of this function, which falls to 0 over a 200–300 ms period following stimulus offset. In each experiment, to-be-reported digit strings were shown for varying durations followed by a noise mask at varying poststimulus intervals. We found the shape of the *performance curve* relating digit-report probability to stimulus exposure duration to be independent of stimulus-mask interstimulus interval. This finding is consistent with the proposition that the iconic-decay function's shape is independent of stimulus duration and allows us to identify this shape. We rejected exponential iconic decay for 6 of 8 observers; however, all observers' decay functions could be adequately fit by *gamma decay*, a generalization of exponential decay.

It has been known for more than 30 years that a briefly presented visual stimulus is generally followed by a fading iconic image (Averbach & Coriell, 1961; Niesser, 1967; Sperling, 1960). The iconic image is often informally represented in a fashion similar to that in Figure 1 (e.g., Averbach & Sperling, 1961). Figure 1a, which depicts the physical stimulus, shows some measure of physical presence (e.g., contrast) as a function of t, the time since stimulus onset. In this example, the stimulus instantly appears, remains on for 100 ms, and then instantly disappears.[1] Figure 1b depicts a psychological correlate of physical stimulus presence. We have labeled the ordinate *proportion available information.* We later define this construct within the context of a specific model; for the moment it is meant to be an intuitive description. The general idea is that all stimulus information becomes available by some time M following stimulus onset, and the information remains available throughout stimulus presence. When the stimulus physically vanishes, available information decays, eventually to zero. This decay of available information constitutes the fading iconic image.

In this article, we report data concerning iconic decay and formulate a descriptive model that accounts for these data. The data, set within the context of the model, allow us to estimate the shape of the iconic-decay function (the right portion of the Figure 1b curve); in particular, we can evaluate the common assumption that such decay is exponential. At the end of the article, we dicuss an alternative class of models that, if viable, would allow generation of the entire $a(t)$ function.

Measuring the Icon's Worth

The paradigm that we used, along with the logic that will carry us from data to conclusions, has its origins in work described by Loftus, Johnson, and Shimamura (1985). Accordingly, we begin with a sketch of this work.

Loftus et al. presented stimuli (naturalistic pictures) for varying durations. The stimuli were followed either by an immediate noise mask or by a noise mask that was delayed by 300 ms following stimulus offset. The noise mask was of much higher luminance and contrast than the stimuli; when the mask and a stimulus were physically superimposed, the stimulus could not be seen at all. Accordingly, it was assumed that no icon followed an immediate-mask stimulus, whereas a relatively complete icon followed a delayed-mask stimulus. Memory for the stimuli was measured in several ways.

Figure 2 shows generic data from this experiment. Here, performance is plotted as a function of exposure duration; we refer to such curves as *performance curves.* As expected, performance increases with exposure duration. The performance advantage for delayed-mask stimuli in relation to immediate-mask stimuli reflects information acquired from the iconic image. Of principal interest is that the delayed- and immediate-mask performance curves were *horizontally parallel,* separated by approximately 100 ms throughout their extent. This finding was quite robust; the constant horizontal curve separation held and was of approximately the same

Portions of the research reported in this article were reported at the 1989 Psychonomic Society Meetings (Atlanta), the 1990 Jackson Hole Interdisciplinary Conference (Jackson Hole, Wyoming), and the 1990 National Research Council conference on visual search (Irvine, California). The research was supported by a National Institute of Mental Health grant to Geoffrey R. Loftus. Raw data in hard copy, MS-DOS, or Macintosh format are available on request.

Very helpful suggestions, critiques, and general guidance on this research project have been provided by Vince Di Lollo, Pete Dixon, Dave Irwin, George Sperling, and Beau Watson. In addition, Di Lollo, Irwin, and an anonymous reviewer provided useful comments on an initial draft of this article.

Correspondence concerning this article should be addressed to Geoffrey R. Loftus, Department of Psychology, University of Washington, Seattle, Washington 98195. Electronic mail may be sent to gloftus @ milton.u.washington.edu.

[1] By "instantly" we mean within a few milliseconds, which might be, for example, the amount of time for a mechanical shutter to open completely or close completely.

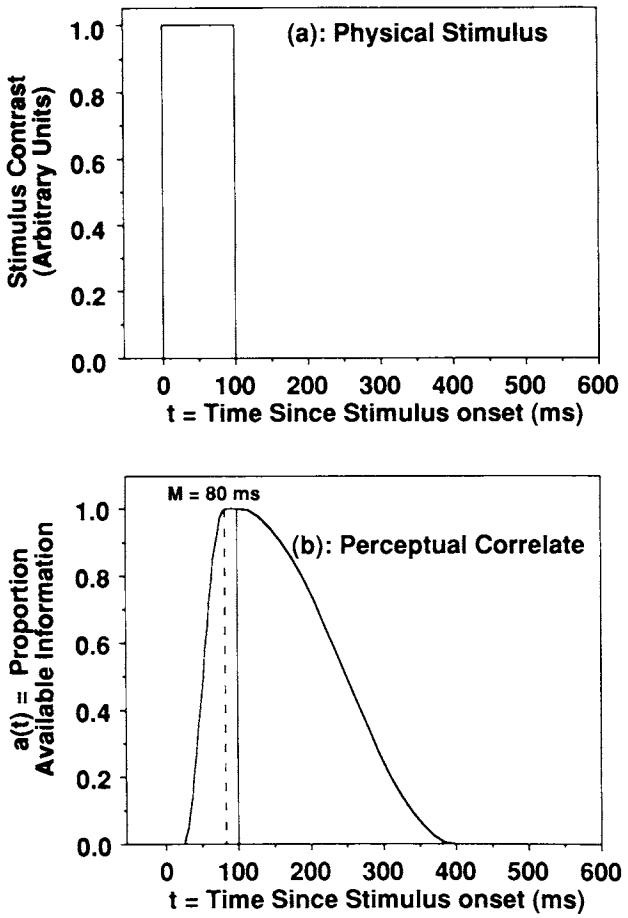

Figure 1. Top panel: Representation of a stimulus input function (contrast as a function of time since stimulus onset). Bottom panel: available information as a function of time since stimulus onset.

magnitude for a variety of memory measures, a variety of stimulus sets, and two stimulus-contrast levels.

At the most general level, one can characterize these results as demonstrating additivity between stimulus duration and stimulus-mask interstimulus interval (ISI). That is, the results could be captured by

$$P(d, 300) = P(d + w, 0), \quad (1)$$

where $P(x, a)$ refers to performance for an x-ms stimulus followed a ms after its offset by a mask, and w refers to the horizontal separation between the curves.

Loftus et al. (1985) interpreted their results in terms of what they called the *icon's worth*, concluding that the information acquired from an icon is the same as the information that can be acquired from an additional w ms of physical exposure duration; that is, the icon was "worth" w ms of additional physical stimulus exposure. As noted, their estimate of w was 100 ms.

The present experiments replicate and extend the Loftus et al. findings. The most important modification is that whereas Loftus et al. used only two stimulus-mask ISIs, the present Experiment 2 used many ISIs. This allowed us to test a more general form of Equation 1:

$$P(d, a) = P(d + w_a, 0). \quad (2)$$

That is, performance engendered by d-ms stimulus followed by an a-ms ISI is equal to performance engendered by a $(d + w_a)$-ms stimulus followed by a 0-ms ISI, where a and w_a are monotonically related. As we shall see, Equation 2 constitutes the prediction of a specific information-acquisition and performance model.

To anticipate, this prediction was confirmed, which in turn allowed us to estimate the shape of the iconic-decay function. We did this by estimating w_a, which we term the icon's *partial worth* at varying times following stimulus offset. These partial worths correspond to the iconic-decay function's integral at varying times, following stimulus offset. As will be described in detail, estimation of this integral provides sufficient information to test various forms of the iconic-decay function itself.

In the remainder of this introduction we describe the experimental paradigm that was used in the present experiments and then describe a model of some psychological processes involved in this paradigm.

The Paradigm

The experimental paradigm is a variant of one that has been used extensively (e.g., Kowler & Sperling, 1980; Rumelhart, 1969; Shibuya & Bundesen, 1988; Sperling, 1963, 1967; Townsend, 1981; Turvey, 1973; van der Heijden, 1978). We have used it in our laboratory and have described it in Loftus

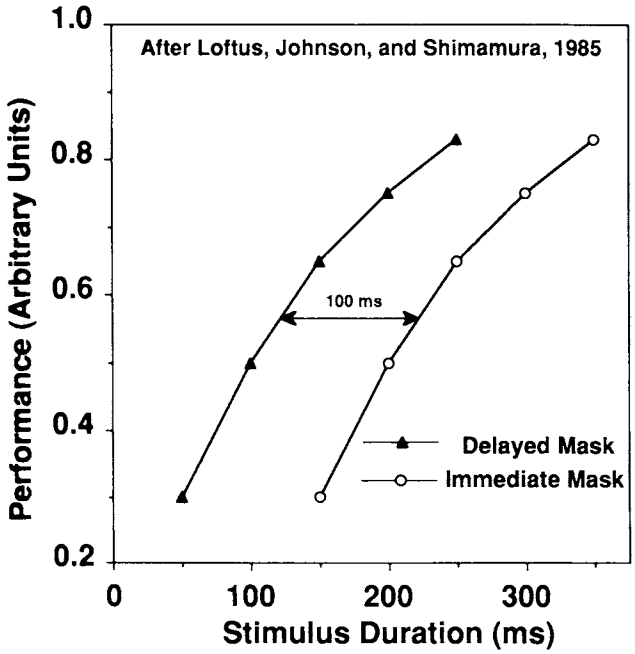

Figure 2. Schematic representation of the Loftus, Johnson, and Shimamura (1985) results. (The delayed- and immediate-mask curves are horizontally parallel, separated by approximately 100 ms.)

(1985), Loftus, Kaufman, Nishimoto, and Ruthruff (in press), and Loftus, Truax, and Nelson (1986).

In each of a series of trials, a stimulus consisting of four digits is presented to an observer whose job is to report immediately as many of the digits as possible in their correct positions. Stimuli are presented briefly (30–300 ms) and are followed by a noise mask that occurs at an a-ms ISI following stimulus offset. The mask's luminance and contrast are such that the mask conforms to Eriksen's (1980) minimal test: When mask and stimulus are visually superimposed, the stimulus cannot be seen at all.

Data From the Paradigm: Performance Curves

The data in this paradigm are captured in a curve that relates the proportion of correctly reported digits (corrected for guessing) to stimulus duration. As noted, we refer to such a curve as a *performance curve.* We have discovered empirically that performance curves in this paradigm can be described essentially perfectly by

$$P = 1.0 - e^{-(d-L)/c} \tag{3}$$

for $d \geq L$ and $P = 0$ otherwise.[2] Here, P is performance, d is stimulus duration, and c and L are free parameters. The parameter L (for "lift-off") represents the minimum duration necessary for above-chance performance, whereas the parameter c determines the curve's shape. Note that c is the duration required for performance to rise from chance to a value of $1.0 - 1/e = 0.63$. Informally, c may be viewed as reflecting the reciprocal of the information-acquisition rate: The greater the information-acquisition rate, the lower is c.

Figure 3 (top panel) shows typical data for 2 observers obtained in our laboratory by using this task. The solid lines represent the best fitting exponential functions. In the bottom panel of Figure 3, the performance measure has been transformed from P to $-\ln(1.0 - P)$. This measure is convenient because Equation 3 implies that it is linear with stimulus duration (which, as is evident in Figure 3, it is). Accordingly, we use it in subsequent descriptions of data from this paradigm.

Present Experiments

In the present experiments, we used factorial designs in which both stimulus-mask ISI and stimulus duration within each ISI level were varied. This design produces a family of performance curves, with one family member corresponding to each ISI level. The major job of the model that we introduce in the next section is to predict the relationship of these performance curves to one another.

A Model of Visual Information Acquisition and Performance

The model that we describe here is a variant of one that has been described elsewhere (Loftus & Hanna, 1989; Loftus, Hanna, & Lester, 1988; Loftus & Hogden, 1988). In what follows we describe the model's assumptions and its application to the present task.

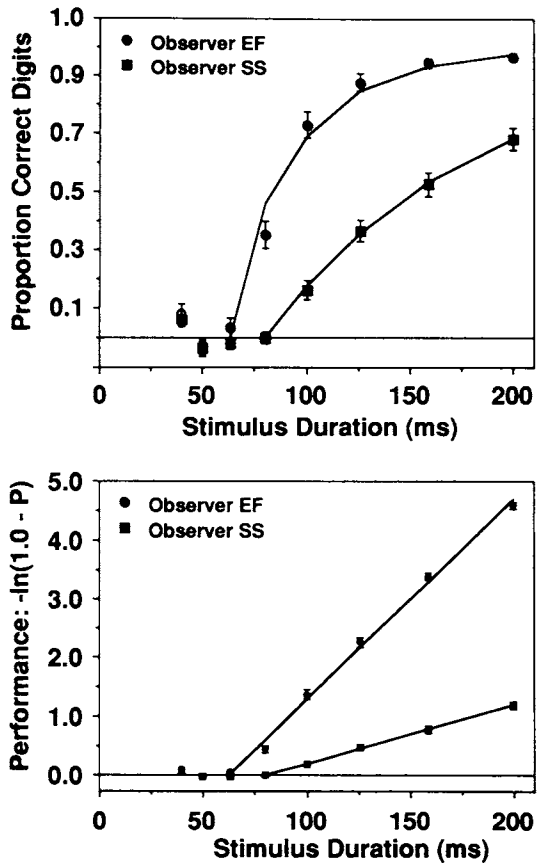

Figure 3. Typical data obtained in the digit-recall paradigm. (Both panels show performance as a function of stimulus duration for two observers. Solid lines show best exponential fits. The performance measure is P, proportion correct in the upper panel and $-\ln[1.0 - P]$ in the bottom panel.)

Assumptions

The model consists of five assumptions, which are as follows.

Available Information

A stimulus consists of information. As indicated in Figure 1b, some proportion of the information is available to be acquired by the observer at any given time, t, following stimulus onset. Available information, $a(t)$, is

$$a(t) = \begin{cases} o(t) & \text{for } t \leq M \\ 1.0 & \text{for } M < t \leq d \\ b(t - d) & \text{for } t > d, \end{cases} \tag{4}$$

where t is time since stimulus onset, d is stimulus duration, M is a free parameter, and o and b are functions. Note that

[2] Variants of this equation have been used to describe the speed–accuracy relationship in speeded-response tasks (e.g., McClelland, 1979; Wickelgren, 1977).

$a(t)$, being the proportion of available information, is a pure number; it has no units.

By Equation 4, all information becomes available at time M following stimulus onset and remains available during the stimulus's physical presence ($t \leq d$). The parameter M does not play an important role in the model; we assume that M is relatively small. Following stimulus offset ($t > d$), available information is described by some function b, which is interpreted to be the iconic-decay function. The only restrictions that we place on b are that it be nonnegative and that its integral from 0 to infinity, $B(\infty)$, be finite. We refer to this integral $B(\infty)$ as w, the icon's *worth*.

Note that because the argument of b is $(t - d)$, the iconic-decay function's shape is, according to the model, independent of the stimulus duration d, as indicated in Figure 4 for a 100- and a 200-ms stimulus. This property seems reasonable and is certainly parsimonious. There is some empirical evidence that iconic decay as measured by partial report is independent of stimulus duration (Irwin & Yeomans, 1986; Yeomans & Irwin, 1985), and this proposition has been explicitly incorporated into iconic-decay models proposed by Irwin and Yeomans as well as by Di Lollo and Dixon (1988). Independence of iconic-decay-function shape and stimulus duration is tested explicitly (and confirmed) in the present Experiment 2.

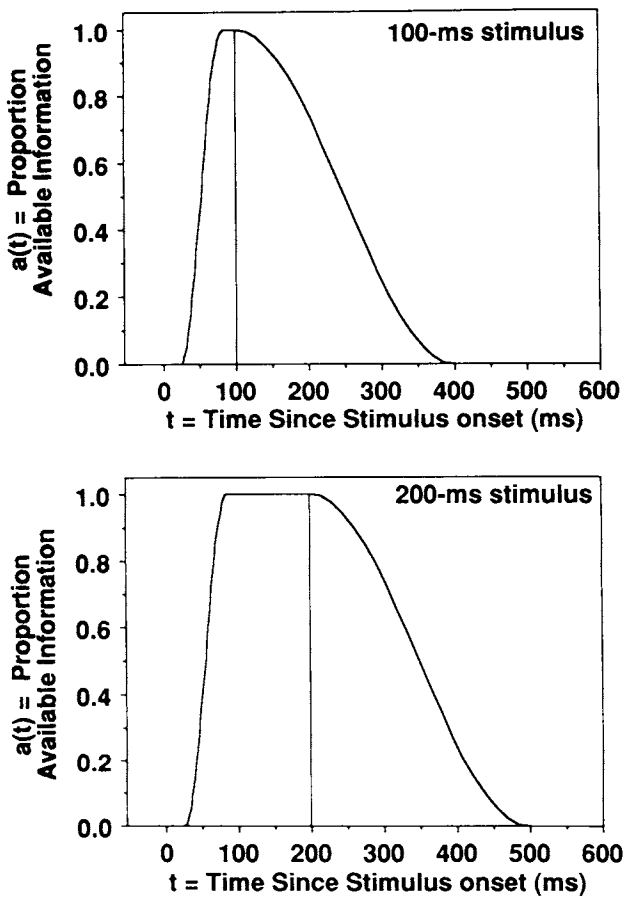

Figure 4. The iconic-decay function of the information-available curve is independent of stimulus duration (stimulus durations of 100 and 20 ms are illustrated).

Information-Acquisition Rate

Available information is acquired by the observer and placed into short-term store, where further processing can continue indefinitely and independently of stimulus presence. At any time t following stimulus onset, some proportion of the available information, $I(t)$, has been acquired. The instantaneous information-acquisition rate at time t, designated $r(t)$, is

$$r(t) = \{a(t)h[I(t)]\}/c. \quad (5)$$

Thus, $r(t)$ is a multiplicative function of $a(t)$, the available information, and some function, h, of I, the proportion of already-acquired information. We assume that the function $h(I)$ is nonnegative, equals 1.0 when $I(t)$ equals 0, and approaches 0 as $I(t)$ approaches 1.0. Finally, c is a scaling constant that has units of time.

This expression for $r(t)$ implies the following properties. First, the smaller is $a(t)$, the available information, the smaller the information-acquisition rate, $r(t)$; in the extreme, when $a(t)$ is zero, $r(t)$ is also zero. Second, because $h(I)$ and thus $r(t)$ approaches 0 as $I(t)$ approaches 1.0, $I(t)$ cannot exceed 1.0.

One parenthetical note is in order. A stronger form of the model would make the stronger assumption that $h(I)$ is strictly monotonically decreasing. This would capture the intuitively reasonable property that the more information that has been acquired from the stimulus, the slower the acquisition of new information. However, this stronger assumption is not necessary for any of the predictions that we make.

Masking

We assume that our mask, which is much brighter and of higher contrast than the stimulus, immediately terminates available information. Thus, if a mask occurs at time $(d + a)$, then $a(t) = 0$ for all $t > (d + a)$.

Unidimensionality of Information

Acquired information, $I(t)$, is unidimensional; that is, it can be represented by a single number on an ordinal scale. At first glance, this assumption seems untenable. Intuitively, it seems that the memory representation of a complex visual stimulus must be multidimensional, and most extant models reflect this intuition. At the very least, such models typically posit item and location information (e.g., Irwin & Yeomans, 1986, Mandler & Parker, 1976).

However, the following interpretation of unidimensionality *is* tenable. Suppose that there are J relevant memory dimensions (e.g., in a model positing item and location information, $J = 2$). At time t following stimulus onset, the memorial representation of the stimulus can be represented by a point in J-dimensional stimulus space. *Encoding* consists of the point's movement along some path through the space, and *information* can be defined as the distance traversed along

this path. By this definition, *information* is unidimensional, whereas the *memory representation* is multidimensional.

Performance

Performance (P) on any memory task (e.g., the digit-recall task described earlier) is monotonically related to amount of acquired information; that is, $P = m(I)$, where m is a monotonic function.

Remark

As noted, we have empirically determined that the function relating performance to stimulus duration in the present task is exponential (see Equation 3 and Figure 3). In Appendix A, we prove that Equation 3 follows if $h[I(t)] = 1.0 - I(t)$ and m is the identity function.

The Model's Application to the Present Paradigm: Estimating the Form of Iconic Decay

In this section we first derive the performance-curve equations and the relations to one another of performance curves corresponding to different stimulus-mask ISI levels. We then show how iconic-decay curve shape can be estimated within the model's context.

Performance Curves for Different ISIs

In Appendix B, we derive an equation for acquired information at time t following the onset of a d-ms stimulus,

$$I(d, t - d) \begin{cases} = H^{-1}[k + d] & \text{for } t \leq d \\ = H^{-1}[k + d + B(t - d)] & \text{for } t > d, \end{cases} \quad (6)$$

where I is information, H^{-1} is a monotonic function, $B(t - d)$ is the integral of $b(t - d)$, the iconic-decay function, and k is a constant.

We refer to $B(t - d)$ as w_a, the icon's partial worth at time $(t - d) = a$ ms following stimulus offset. Accordingly, from Equation 6 we can infer that information acquired from a d-ms delayed-mask stimulus ($a > 0$) is

$$I(d, a) = H^{-1}[k + d + w_a]. \quad (7)$$

From Equation 6, we can also infer that information acquired from an immediately masked $(d + w_a)$-ms stimulus is

$$I(d + w_a, 0) = H^{-1}[k + (d + w_a)]. \quad (8)$$

Because the right sides of Equations 7 and 8 are the same, the left sides are the same as well, or

$$I(d, a) = I(d + w_a, 0). \quad (9)$$

Finally, because performance, P, is assumed to be monotonically related to information, I, two conditions producing the same amount of information also must produce equal performance. This allows us to write the equation for our fundamental prediction:

$$P(d, a) = P(d + w_a, 0). \quad (10)$$

The prediction embodied in Equation 10 is thus: Performance engendered by a d-ms stimulus followed by an a-ms ISI is the same as performance engendered by a $(d + w_a)$-ms stimulus followed by a 0-ms ISI, where w_a is independent of stimulus duration d. This implies that a factorial design in which d and a are varied should result in a family of horizontally parallel performance curves, one curve for each value of a, with the a-ms ISI curve separated from the 0-ms ISI curve by w_a ms. As depicted in Figure 2, we have seen that this prediction was confirmed by Loftus et al. (1985) for an a value of 300 ms.

The Iconic-Decay Function's Shape

Consider the kind of stimulus-duration–stimulus-mask ISI factorial experiment that we have been describing. Confirmation of Equation 10 in such an experiment is consistent with the proposition that iconic-decay-function shape is independent of stimulus duration. Accordingly, such confirmation permits us to proceed under the supposition that there is only a single iconic-decay shape to measure, as illustrated in Figure 4. In what follows, we describe how this putative single shape is estimated under such circumstances.

The horizontal distance between an a-ms ISI performance curve and the zero-ISI performance curve constitutes an estimate of $w_a = B(t - d)$, the integral of the iconic-decay function from $t = d$ to $t = (d + a)$. Accordingly, the data from this paradigm permit us to estimate directly the shape of $B(t - d)$.

Ideally, we could then recover the original iconic-decay function, $b(t - d)$, from its integral. Unfortunately, this is not possible because any real experiment, measuring performance at a discrete set of ISIs, allows one to estimate the value of $B(t - d)$ only at discrete times (viz. the particular ISIs used in the experiment), and there is a one-to-many relationship between an integral, $B(x)$, evaluated at discrete points, and its corresponding derivative, $b(x)$.

Accordingly, to estimate the decay function's shape, we must nominate *candidate functions* (e.g., exponential decay is a reasonable candidate function) and compare the candidate-function integrals with the empirically estimated integrals. The adequacy of any candidate function's fit can then be evaluated with standard statistical procedures.

Visible Persistence

Although it is not the focus of this article, we wish to comment on the model's application to visible persistence. Several researchers, particularly Coltheart (1980), have elucidated difficulties in explaining both informational and visible persistence within the same model. The crux of the problem is that two important variables—stimulus duration and stimulus luminance—have a null or a positive effect on information acquisition from the icon (as measured by partial report) but strong negative effects on visible-persistence duration (cf. Adelson & Jonides, 1980; Bowen, Pola, & Matin, 1974; Di Lollo & Dixon, 1988; Efron, 1970a, 1970b; Haber & Standing, 1970).

Available Information and Information-Acquisition Rate

Our model posits that the iconic-decay function, $b(t)$, is independent of stimulus duration. Accordingly, if visible persistence were posited to be determined by $b(t)$, the model's prediction about the stimulus-duration effect on visible persistence would be incorrect. However, for a variety of reasons, we have assumed visible persistence at time t to be determined by the *information-acquisition rate*, $r(t)$. The key to this assumption's success in explaining visible persistence is that $r(t)$ declines with acquired information, I; accordingly it is lower at the offset of both longer duration and higher luminance stimuli. Elsewhere, Loftus and his colleagues have shown how this assumption quantitatively accounts for both duration and luminance effects in temporal-integration and synchrony-judgment tasks (Loftus & Hanna, 1989; Loftus & Hogden, 1988).

Relation to Existing Models

This account of the information–visible-persistence relationship is formally similar (although conceptually dissimilar) to accounts proposed by Irwin and Yeomans (1986) and Di Lollo and Dixon (1988). In these models, there is a visual (i.e., an analogue or schematic but *nonvisible*) representation that begins to decay following stimulus offset (such decay is explicitly exponential in Di Lollo and Dixon's model) and is independent of stimulus duration. In addition, in Di Lollo and Dixon's model there is a separate visible representation that begins to decay at stimulus onset (thus accounting for the inverse effect of stimulus duration on tasks requiring visible persistence).

Formally, the present model's poststimulus available information, $b(t)$, corresponds to the nonvisible representation in both the Irwin and Yeomans and Di Lollo and Dixon models, whereas the present model's information-acquisition rate, $r(t)$, corresponds to Di Lollo and Dixon's visible representation.

EXPERIMENTS

We report two experiments. Experiment 1 was designed to replicate the Loftus et al. experiments with simpler stimuli, a more powerful experimental design, and a more sophisticated statistical assessment technique. Experiment 1 permits a test of the Equation 10 prediction with two ISI values.

Experiment 2, an extension of Experiment 1, was designed to test the general Equation 10 prediction when many ISI values are used. Assuming the prediction's confirmation, the Experiment 2 data further permit us to estimate the iconic-decay function's shape by using the technique that we have just described.

Experiment 1

As noted, Loftus et al. (1985) found immediate- and delayed-mask performance curves to be horizontally parallel (see Figure 2). This finding suggests that information acquired from an icon is equivalent to information acquired from a w-ms extension of the physical stimulus, where $w = B(\infty)$ is independent of stimulus duration. Such a conclusion, if valid, would be interesting for several reasons. First, it provides a simple, parsimonious description of information acquisition from an icon. More generally, it constrains models that are designed to account for such information acquisition. In particular, it strongly favors models (like the one described earlier) in which the nature of information acquired from the icon is qualitatively the same as information acquired from the physical stimulus.

There were, however, two shortcomings in the Loftus et al. procedure. First, their experiments involved gathering a relatively small amount of data from each of many observers, which means that (1) subtle averaging artifacts could have obscured the main results, and (2) icon-worth estimates could not be obtained from individual observers. Second, Loftus et al. did not use any formal statistical procedure to evaluate the hypothesis that immediate- and delayed-mask performance curves are horizontally parallel.

Experiment 1 was designed to replicate the Loftus et al. results with these shortcomings removed. Two major procedural modifications were introduced. First, a great deal of data were collected from individual observers. Second, a more sophisticated statistical procedure was used to assess the hypothesis that immediate- and delayed-mask performance curves are parallel. This procedure rests on the finding that Equation 3 provides an adequate description of performance curves. Because the parameter c determines a performance curve's shape, different performance curves are horizontally parallel if and only if the c values for the curves are identical (see Appendix C). Identical c values constitute a testable null hypothesis.

Method

The basic experimental procedure consisted of a series of trials. On each trial, a four-digit string was presented for some exposure duration (on the order of 30–400 ms) followed by a mask at either a 0- or 250-ms ISI. The observer's task was to report immediately as many of the digits as possible in their correct positions, guessing if necessary.

Observers

Four observers participated in the experiment: the 3 authors and an undergraduate female observer. All observers were intimately familiar with the purposes of the experiments. All observers were highly practiced, having participated in a minimum of 3,000 practice trials prior to beginning the experiments.

Stimuli and Apparatus

Observers viewed stimuli and a random-noise mask, all prepared as 35-mm slides. A stimulus consisted of a 4 (columns) × 3 (rows) array of black digits on a white background. Eighty such stimuli were prepared and used repeatedly. The $4 \times 3 \times 80 = 960$ digits composing all stimuli were selected randomly and with replacement from the set of 10 digits. Each digit subtended a visual angle of 0.56° vertically and 0.28° horizontally. Digits were separated by 0.37° vertically and 0.74° horizontally. On a given experimental trial, one four-digit row of one stimulus was the to-be-reported target. Target row was blocked

over trials; accordingly, an observer always knew far in advance which row was the target.[3] The noise mask consisted of black visual noise on a white background.

Stimulus contrast was substantially reduced to avoid ceiling performance. Contrast reduction was accomplished by (1) attenuating stimulus luminance with a neutral-density filter and (2) by superimposing a uniform adapting field over the stimuli. Stimulus contrast varied from observer to observer. A summary of luminances and contrasts for both experiments is provided in Table 1. Note that in Table 1 observers are referred to by their initials.

All stimuli were displayed by using Kodak projectors equipped with Gerbrands tachistoscopic shutters. A random-access projector was used to display the stimuli, whereas standard carousel projectors were used to present the noise mask, the contrast-reducing uniform light field, and a fixation point that initiated each trial. Responses were made on a numeric keypad. All display equipment was enclosed in a soundproof box. All display and response collection was under the control of an AT-compatible computer system described by Stoddard and Loftus (1988).

Design and Procedure

Twelve conditions were defined by two ISI levels (0 and 250 ms) and six exposure durations within each ISI level. The selection of 250 ms as the longest ISI resulted from pilot work which indicated that digit-recall performance, though improving as ISI was increased from 0 to 250 ms, did not improve further as ISI was increased beyond 250 ms. The exposure-duration values within each ISI level were selected with the goals of (1) producing roughly equal performance ranges within each ISI level (which meant that durations within the 0-ms ISI level had to be suitably greater than corresponding durations within the 250-ms ISI level) and (2) maintaining performance (proportion of correctly recalled digits) within a performance range of roughly 0.1–0.9 (proportion correct). These goals were accomplished by trial and error over a great deal of pilot work.

For each observer, the exposure durations were specified by three experimental parameters. *Base* was the minimum duration used in the 250-ms ISI level, *ink* was the stimulus-duration increment from each duration level to the next duration level within each ISI level, and *worth* was a duration added to each 250-ms ISI duration to produce a corresponding 0-ms ISI duration.[4] Table 2 provides an example of the 12 exposure durations generated by the following parameter set: base = 30 ms, ink = 40 ms, and worth = 90 ms.

Table 1
Luminances and Contrasts: Stimulus Contrasts for Different Observers (Represented by Initials) Are Shown for Experiments 1 and 2

Stimulus/observer	Background luminance	Foreground luminance	Contrast
Experiments 1 and 2			
Dark background field	0.34	—	—
Fixation point	0.46	1.79	0.591
Adapting field	23.73	—	—
Mask	67.17	2.35	0.933
Experiment 1			
JD, CG, GL	33.18	24.72	0.146
PG	36.75	26.46	0.163
Experiment 2			
PG, GL	27.51	24.49	0.058
CA, TB, JD, CG, LK	29.40	24.72	0.086
SO	33.18	24.72	0.146

Note. Luminances are in candles/m^2.

Table 2
Illustration of Exposure-Duration Composition (in Milliseconds) for Experiment 1

Duration	Interstimulus interval	
	0 ms	250 ms
1	120	30
2	160	70
3	200	110
4	240	150
5	280	190
6	320	230

Note. All durations assume a 30-ms base, 40-ms ink, and a 90-ms worth.

Finally, a fourth experimental parameter, stimulus contrast, completed the specification of the experimental configuration for each observer. The complete configurations for all observers are summarized in Tables 1 and 3.

Each observer participated in 18 blocks of 80 trials per block. Recall that stimuli were prepared as 3 four-digit rows. On any given block, only one row (top, middle, or bottom) was the to-be-reported target.

The event sequence for a given 80-trial block was as follows. First, a high, medium, or low tone (2000, 1000, or 500 Hz) signaled the observer that the top, middle, or bottom row would be the target row for that block (i.e., for the next 80 trials). Next, eight practice trials were presented. The conditions for these practice trials were selected randomly and without replacement. Next, 72 experimental trials were presented. The 12 conditions were randomly intermingled over the 72 trials, with the restriction that each condition occurred exactly twice within each 24-trial sequence. Stimulus-presentation order was quasirandom.[5]

[3] A stimulus slide consisted of 3 four-digit rows rather than just 1 for two reasons. The first, less important reason is historical: We had done some partial report work in our laboratory, and the stimulus slides that we had used were conveniently hanging around. The second, more important reason is that we wanted to use the same slides repeatedly over observers. It was easier to do this (i.e., there was less inclination to memorize the digit sequences) with three digit rows per slide because an observer only saw a given row as a target once every three blocks.

[4] Note here that we are using "worth" to denote an experimental rather than a theoretical parameter. The experimental and theoretical uses of the term are related in that if a given observer's empirical and theoretical worths were identical, then horizontally parallel immediate- and delayed mask performance curves would cover identical performance ranges for that observer. The experimental value of worth that we selected for each observer constituted our best guess, based on pilot work, about what an observer's theoretical worth would turn out to be.

[5] The 80 stimulus slides were fixed in the 80 slots of a carousel tray. We wanted to preclude the observers' ability to memorize and make use of sequential slide-to-slide information (e.g., we did not want Middle Row 7184 to always follow Middle Row 0072). We accomplished this goal as follows. On each block, the 72 experimental stimuli were randomly divided into two 36-stimulus groups (Groups A and B). The carousel circled twice within each block: On Pass 1, all Group A stimuli were shown, and on Pass 2, all Group B stimuli were shown. This scheme ensured that stimulus ordering differed unpredictably from one block to the next.

Table 3
Experimental Parameters for Experiment 1

Observer	Base	Ink	Worth
GL	40	30	100
PG	80	30	80
JD	70	20	70
CG	30	30	95

The to-be-reported row was changed systematically over blocks (in a top–middle–bottom–top . . . sequence; thus, each row served as target in 6 of the 18 blocks). Assignment of conditions to trials within a block was also changed over blocks.

As noted, a block consisted of 80 trials. The event sequence within each trial was as follows. First, there was a 500-ms fixation point accompanied by a warning tone. Warning-tone frequency was 2000, 1000, or 500 Hz and reminded the observer which row (top, middle, or bottom) was the target during the current block. The fixation point was always positioned in the middle of the upcoming slide, that is, between the second and third digits of Row 2. Following the fixation point/warning tone was the stimulus, presented for its appropriate duration, followed by a dark 0- or 250-ms ISI, followed by the mask, which was presented for 300 ms. The mask was followed by the adapting field, which remained present for the rest of the trial. The observer typed in four responses after mask disappearance, guessing on a digit if uncertain. Following responding, there was feedback in the form of four 150-ms beeps. Each beep was 2000 Hz if the corresponding digit had been correctly reported and 500 Hz if the corresponding digit had not been correctly reported. Following feedback was a 300-ms dark interval prior to the start of the next trial.

Results

We had two main goals in data analysis. First, we wanted to determine whether the delayed- and immediate-mask performance curves were horizontally parallel as predicted by the model. Second, given the model's confirmation, we wanted to measure the horizontal difference between the curves; this distance constitutes an estimate of the icon's worth.

Figure 5 shows $-\ln(1.0 - P)$ as functions of stimulus duration for each of the two ISIs.[6] For each observer the solid triangles represent delayed-mask conditions, and the open circles represent immediate-mask conditions. The solid lines through the data points are the best linear fits constrained to have equal slopes (and thus equal c values) for the two curves. (Note that both the duration and performance axes vary across the 4 observers.) In what follows, we refer both to Figure 5 and to Table 4, which summarizes the Experiment 1 data.

Evaluating the Horizontal-Parallelism Prediction

In Table 4, the two rows under "Separate c-value fits" provide the c values for each of the 4 observers, calculated independently for the two ISI levels.[7] The mean values are virtually identical and are not significantly different from one another.

Our model predicts horizontally parallel immediate- and delayed-mask performance curves. Given validity of the linear fit, horizontally parallel performance curves imply and are implied by equal c values for the immediate- and delayed-mask curves, as shown in Appendix C. This prediction can be embodied as the following null hypothesis,

$$-\ln(1.0 - P_{ij}) = (d_{ij} - L_j)/c \qquad (11)$$

where j indexes the stimulus-mask ISI level (0 or 250 ms in this experiment), and i indexes stimulus duration within each ISI level. Note that separate L_js are allowed for the two ISI levels, but c remains constant over ISI.

To test Equation 11's adequacy, we used standard linear regression procedures to estimate three parameters for each observer: two lift-offs (L_0 and L_1 for immediate- and delayed-mask curves) and a common slope, c. The best fitting parameter values are shown in Table 4: Rows 1–3 under "Common c-value fits."

The Pearson r^2 values resulting from this common-c fitting procedure are shown in Row 5 ("Obtained r^2") under "Common c-value fits." Note that each r^2 is based on 9 degrees of freedom: 12 data points minus 3 estimated parameters.

The fits are quite good: The lowest r^2 is .981. To assess the statistical validity of the model, we carried out a 1,000-iteration computer simulation for each observer. On each iteration, we generated the 12 predicted condition means by using that observer's best fitting parameters and then perturbed the simulated mean for each condition by an amount drawn randomly from a normal distribution with a mean of zero and a standard deviation equal to that condition's standard error. We computed the r^2 for each iteration's data set, thereby obtaining the r^2 distribution given the model's validity and the observed error variance. The r^2 values corresponding to the upper and lower 5% of these distributions are provided in Table 4: Rows 4 and 6 under "Common c-value fits." The obtained r^2s fall within or above this range. In short, the model is statistically confirmed.[8]

Individual Differences in the Icon's Worth

Loftus et al. (1985) estimated the icon's worth, w, to be about 100 ms. As noted, this value emerged consistently from a variety of experimental conditions. However, in all cases, it was estimated from data that had been averaged over large numbers of observers. The present data allow us to estimate the icon's worth for each individual observer. These estimates (which are simply the difference between the L values for the

[6] Recall that P is the proportion of correct digits corrected for guessing (guessing probability = 10%). The correction formula was $P = (p - 0.1)/0.9$, where p is the raw proportion, and P is the corrected proportion.

[7] The c values are similar for 3 of the 4 observers, but c is substantially greater for GL. GL, age 42, was the oldest observer; PG, JD, and CG were in their early 20s. Loftus, Truax, and Nelson (1986) showed strong age effects on the information-acquisition rate in this task; however, in that study, 20-year-olds were compared with 70-year-olds. It appears (alas) that age deterioration may begin earlier.

[8] By "the model" we mean the strong version of the model, in which the function h is linear and m is the identity function. Confirmation of the stronger version ipso facto confirms the weaker version, in which h and m are less constrained.

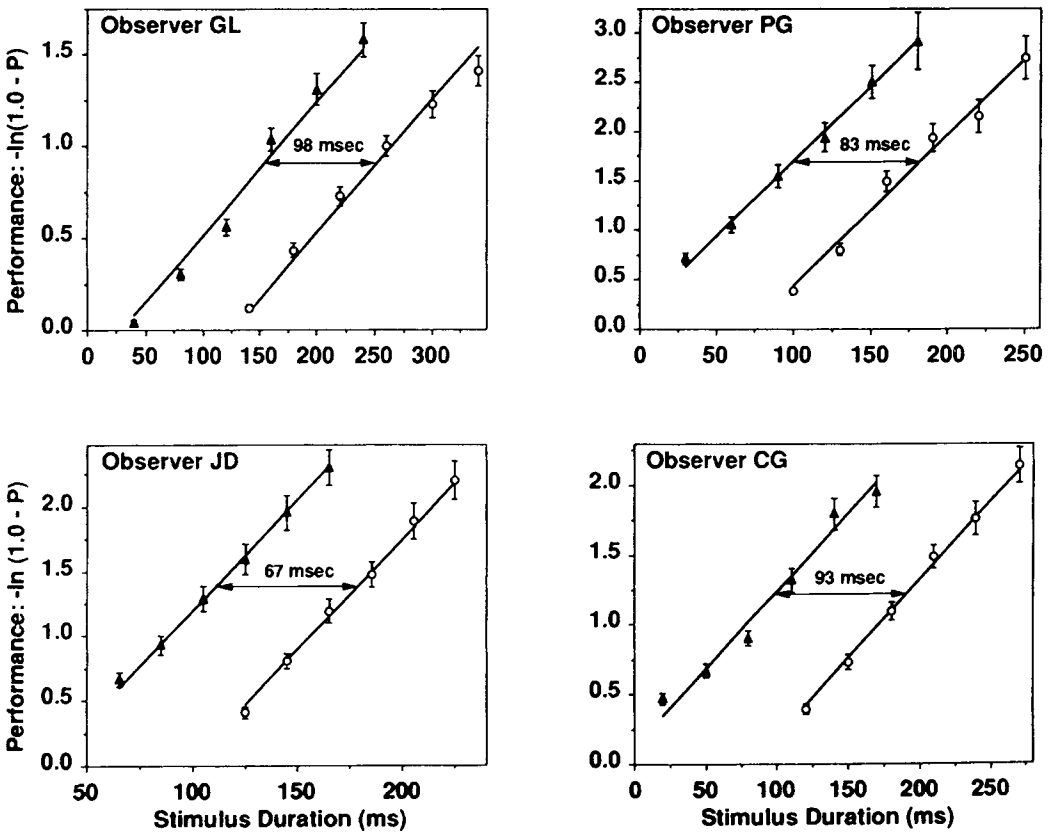

Figure 5. Results of Experiment 1: Performance curves for individual observers. (Closed triangles represent data from the 250-ms interstimulus interval (ISI) levels, and open circles represent data from the 0-ms ISI levels. Solid lines are best fitting linear functions constrained such that within each panel the two curves have the same slope.)

0- and 250-ms ISI levels) are provided in Table 4 under "Icon worth." We have also provided standard errors for these estimated worths, which were obtained from the same computer simulation described earlier. There are clear differences among the 4 observers; the estimate of w ranges from 68 ms for JD to 98 ms for GL. It is not our goal to explain these individual differences here; we merely note them for the record.

Table 4
Summary Data (in Milliseconds) for Experiment 1

	Observer				
Fits/icon worth	GL	PG	JD	CG	*M*
Separate c-value fits					
0 ms	154	81	56	86	94
250 ms	125	85	61	94	91
Common c value fits					
Common c	139	83	58	90	
L_0	112	71	92	82	
L_1	14	−12	24	−11	
5th percentile r^2	.977	.942	.951	.966	
Obtained r^2	.981	.993	.997	.987	
95th percentile r^2	.997	.992	.992	.994	
Icon worth: $L_0 - L_1$					
Worth	98	83	67	93	
SE	4.7	5.5	3.5	4.4	

Discussion

The conclusions to be drawn from the Experiment 1 results are straightforward. First, the horizontal-parallelism prediction of our model is confirmed. This replicates the Loftus et al. findings with a more powerful experimental paradigm and a more sophisticated statistical-analysis technique. Second, we have obtained individual estimates of the icon's worth for 4 observers.

Experiment 2

We had two major goals in Experiment 2. The first was to assess Equation 10's validity when many ISIs were used. Confirmation of the Equation 10 prediction is consistent with the hypothesis that iconic-decay-function shape is invariant across *d*, the stimulus duration.

To anticipate, the Equation 10 prediction was confirmed. Accordingly, our second goal was to determine the iconic-decay function's shape. As described earlier, we must do this by nominating and evaluating candidate functions. Our null candidate was *exponential decay*, which we selected for several reasons. First, because exponential decay arises from an environmentally common (independent Poisson) process, many physical entities decay exponentially (e.g., radioactive decay is exponential). Second, exponential decay seems to be a default assumption in the literature for many information-loss processes (e.g., Bogartz, 1990; Murdock & Cook, 1960), including iconic decay (e.g., Di Lollo, 1984; Di Lollo & Dixon, 1988; Hawkins & Shulman, 1979).

Method

The Experiment 2 method was similar to the Experiment 1 method except that we used more than two ISIs. The stimuli, apparatus, task, and general methodology were identical to Experiment 1. The observers were 4 original Experiment 1 observers, GL, PG, JD, and CG, plus 4 new observers, TB, SO, CA, and LK. Table 1 includes stimulus luminances and contrasts for all observers.

The Experiment 2 design included either six ISIs and six stimulus durations within each ISI level, for 36 total conditions (original observers), or seven ISIs and five stimulus durations within each ISI level, for 35 total conditions (new observers). The old observers had 8 practice trials followed by 72 experimental trials per block, whereas the new observers had 10 practice trials followed by 70 experimental trials per block. For all observers, there were two instances of each condition per block. For reasons to be described later, one of the original observers (CG) participated in both the 6 × 6 and the 7 × 5 configurations. The number of blocks per observer were 108 (JD and CG in the 6 × 6 configuration), 73 (GL), 65 (PG), 70 (CA, SO, and LK), 41 (TB), and 35 (CG, 7 × 5 configuration).

The ISI and stimulus-duration values for a given observer were generated as follows. The extreme ISIs were 0 and 250 ms as in Experiment 1. The intermediate ISIs were selected to provide approximately equal icon-worth increments across the ISI levels. This selection was based on the Experiment 1 data, pilot work, and our guesses about the iconic-decay function's shape. As in Experiment 1, stimulus durations within each ISI were selected with the goal of achieving approximately equal performance ranges within each ISI level.

Results

The results were similar across the 8 observers. For expositional simplicity, we first describe the complete data-analysis process for a single observer and then summarize the findings for all 8 observers.

Analysis of a Single Observer

Figure 6 shows complete data from 1 observer (JD). As in Figure 5, the solid triangles and open circles represent the 250-ms and 0-ms ISI levels, respectively. The intermediate ISI levels are represented, from longer to shorter ISIs, by solid squares, solid circles, open triangles, and open squares. The six solid curves for JD's six ISI levels were obtained by fitting Equation 11 through standard regression techniques. As in Experiment 1, we constrained *c*, the curve's shape parameter, to be identical for all six curves, and allowed only *L*, the *d*-intercept parameter, to vary. Even with this constraint, JD's fit is very good; the overall Pearson r^2 is .986. The numbers at the right of the curves indicate the five nonzero ISI durations (number before the slash) and the horizontal separation between the corresponding performance curve and the zero-ISI (rightmost) performance curve (number after the slash).

Evaluation of Equation 11's adequacy. We assessed Equation 11's adequacy by using the same computer-simulation technique as in Experiment 1. Again, we ran a 1,000-iteration simulation. On each iteration we generated a simulated data set by perturbing each of JD's 36 predicted condition means by an amount equal to that condition's observed standard error. This procedure provided the distribution of Pearson r^2s and the standard error of each partial-worth estimate, given Equation 11's validity and the observed error variance in the data. For JD, the observed $r^2 = .986$ fell at the 92nd percentile of the simulated distribution. This means that we fail to reject

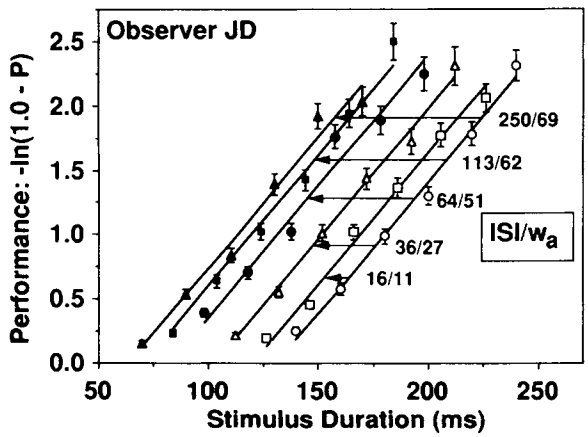

Figure 6. Results of Experiment 2 for Observer JD: Performance curves for six interstimulus interval (ISI) levels. (Solid curves are best fitting linear functions constrained such that all curves have the same slope. Numbers at the right represent ISI value [number before the slash] and horizontal distance to the 0-ms ISI [rightmost] curve [number after the slash].)

the null hypothesis embodied in Equation 11 at the $p = .92$ level.

Independence of the iconic-decay function's shape and stimulus duration. Recall our hypothesis, illustrated in Figure 4, that the iconic-decay function's shape is independent of stimulus duration. As we demonstrated earlier (see Equation 10 and accompanying text), this hypothesis predicts performance curves for different ISIs to be mutually parallel. Accordingly, our failure to reject horizontal parallelness confirms this hypothesis for JD.

Evaluating the exponential-decay hypothesis. Figure 7a shows JD's estimated partial icon worths (i.e., the estimate of JD's $B[t - d]$ values) as a function of stimulus-mask ISI. As indicated, these data points are simply the best fitting horizontal differences between the Figure 6 zero-delay performance curve and each of the other five curves. As described earlier, the standard errors resulted from the computer simulation.

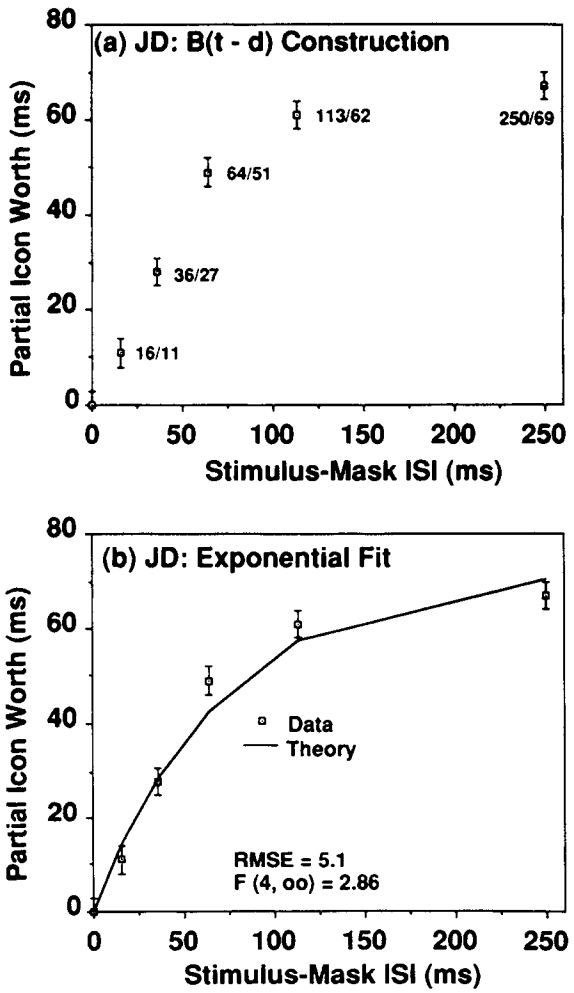

Figure 7. Top panel: Generation of JD's $B(t - d)$ function. (Partial icon worth values [obtained from Figure 6] as a function of interstimulus interval. Numbers represent the same thing they did in Figure 6.) Bottom panel: JD's best fitting exponential-decay function for her top-panel data.

The exponential-decay hypothesis is denoted by

$$b(t - d) = e^{-(t-d)/w}, \tag{12}$$

where b is the iconic-decay function, $(t - d)$ is ISI, and w is the total area under the iconic-decay curve (the icon's worth). Integrating $b(t - d)$ from 0 to $(t - d)$,

$$B(t - d) = w_a = w[1.0 - e^{-(t-d)/w}], \tag{13}$$

where w_a is the icon's partial worth at time $a = (t - d)$ ms following stimulus offset.

We fit Equation 13 to the Figure 7 data, allowing one free parameter, w. The best fitting exponential-decay prediction is shown as the solid line in Figure 7b. The prediction is systematically discordant with the data. The root-mean-square error (RMSE) is 5.1 ms, and the exponential fit can be statistically rejected, $F(4, \infty) = 2.86$. JD's iconic-decay function, in short, is not exponential.

Analysis of All 8 Observers

Table 5, which provides summary data for all 8 Experiment 2 observers, is organized as follows. The 4 male observers are shown first, followed by the 4 female observers. As noted, CG, a female observer, participated in both the 6 (ISIs) × 6 (durations) experimental configuration and the 7 × 5 configuration. The data from these two configurations are referred to as *old* and *new*.

Evaluation of Equation 11's adequacy. The first two rows under "Equation 11 fit" in Table 5 show the (common) c value along with the L value for the 0-ms ISI level.

Row 4 under "Equation 11 fit" shows the observed Pearson r^2 (based on 29 and 27 degrees of freedom for the 6 × 6 and 7 × 5 configurations). Rows 3 and 5 show the 5th and 95th percentiles of the r^2 distribution obtained from the computer simulations. Seven of the observed r^2s fall within the 5%–95% range. The two r^2s that are slightly less than the 5th percentile are still high (.980 and .966). In short, the fits are good.

Estimation of partial worths. The first six rows under "ISI/icon worth" in Table 5 show partial icon worths for the five or six nonzero ISIs. All cells in this table have two entries: The number preceding the slash is the stimulus-mask ISI, and the number following the slash is the estimated partial icon worth. The seventh row shows the standard error of the estimated partial worths obtained from the computer simulations.

Evaluation of the exponential-decay hypothesis. The three rows under "Exponential-decay function fits" in Table 5 show the best fitting w value (Row 1), and RMSE (Row 2), and the F ratio (Row 3). Note that exponential decay can be rejected for 6 of the 8 observers. One observer (CA) had a fairly large RMSE, however; her partial-worth estimates were too noisy (SE = 6.5 ms) to be able to reject the exponential-decay hypothesis (or any other reasonable hypothesis).

Figure 8 provides a graphical representation of the exponential fit for all observers. To prepare Figure 8, we *normalized* each observer's data as follows. We chose an (arbitrary) value of 100 ms for w, the icon's worth (also known as the

Table 5
Summary Data for Experiment 2 and Gamma-Decay Constraining in Values (in Milliseconds)

	Observer								
	Male				Female				
Fits/icon worths	GL	PG	SO	TB	JD	LK	CA	CG (old)	CG (new)
				Summary data					
Equation 11									
c	116	211	76	75	51	97	72	72	88
L_0	111	142	36	143	111	150	125	98	104
5th percentile r^2	.980	.973	.940	.966	.965	.982	.957	.983	.960
Obtained r^2	.985	.966	.964	.970	.986	.983	.987	.980	.966
95th percentile r^2	.992	.990	.978	.988	.988	.994	.991	.994	.989
ISI/icon worth									
Worth 1	22/21	18/8	14/13	15/14	16/11	13/9	15/11	21/12	16/13
Worth 2	51/50	41/38	29/35	31/27	36/27	26/22	31/28	49/38	33/27
Worth 3	92/74	73/63	46/41	48/44	64/51	40/37	48/43	87/58	51/35
Worth 4	161/97	129/78	66/52	69/62	113/62	58/49	69/51	153/72	73/52
Worth 5	250/102	250/87	94/63	99/75	250/69	83/59	99/66	250/80	105/63
Worth 6	—	—	250/75	250/92	—	250/77	250/80	—	250/75
SE	3.5	4.3	4.8	5.2	3.0	3.3	6.5	3.3	3.7
Exponential-decay function									
w	126	105	84	109	73	85	89	86	81
RMSE	8.6	10.3	7.8	8.4	5.1	5.7	5.2	4.4	3.5
F ratio	6.09	5.66	2.61	2.59	2.86	3.06	0.62	1.80	0.90
				Gamma-decay constraining n values					
$n = 2$									
RMSE	3.4	6.6	4.2	3.7	3.5	2.6	2.5	5.6	4.3
F ratio	0.95	2.30	0.75	0.50	1.37	0.60	0.14	2.97	1.42
$n = 3$									
RMSE	1.8	5.5	3.8	2.1	4.3	3.0	3.7	7.2	6.2
F ratio	0.26	1.63	0.62	0.16	2.09	0.83	0.32	4.93	2.88
$n = 4$									
RMSE	2.3	5.4	4.2	2.3	5.1	3.9	4.8	8.3	7.3
F ratio	0.45	1.54	0.75	0.20	2.84	1.43	0.54	4.02	6.59
$n = 5$									
RMSE	3.0	4.3	4.7	2.9	5.7	4.6	5.6	9.0	8.1
F ratio	3.57	1.76	0.97	0.31	3.57	1.97	0.74	7.69	4.97

Note. Observers GL, PG, and JD are original observers (six interstimulus intervals [ISIs]). Observers SO, TB, LK, and CA are new observers (seven ISIs). Observer CG served as both an original and a new observer. For the ISI/icon-worth data, the number to the left of the slash is stimulus-mask ISI; the number to the right of the slash is estimated icon worth. *SE*s apply to estimated icon worths. RMSE denotes root-mean-square error.

exponential iconic-decay constant; cf. Equation 13). We then computed a normalization factor equal to $100/w_i$, where w_i is the best fitting estimate of the ith observer's decay constant. For observer i, we then multiplied ISIs, observed partial worths, and standard errors by w_i, thereby rendering all observers' functions comparable. The results of this analysis are summarized in Figure 8, wherein all normalized partial worths are plotted as a function of normalized ISI. The best fitting exponential function departs systematically and significantly from the observed function.

Observer CG

As indicated in Table 5, CG was the only observer for whom the exponential-decay function fit adequately. It was to verify this finding that CG ran through a second set of blocks in the 7 × 5 configuration. CG's data are consistent with exponential decay in both configurations.

Discussion

We first discuss our implications of the good Equation 11 fit to the data, and second we discuss our rejection of exponential iconic decay.

Equation 11's Fit to Performance Curves

All observers showed a good fit to the Equation 11 performance-curve descriptions. This finding has two implications. First, Equation 11's fitting correctly confirms that performance curves from all ISI levels are horizontally parallel to one another. This finding allows the conclusion that the iconic-decay function's shape is independent of physical-stimulus duration; such a conclusion is in accordance with the theory and data of Di Lollo and Dixon (1988), Irwin and Yeomans (1986), and Yeomans and Irwin (1985).

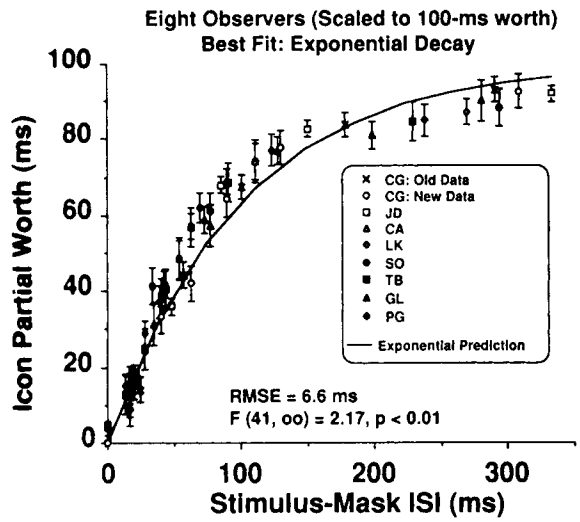

Figure 8. Best fitting exponential-decay function for 8 observers. (Observers' data have been scaled [as described in the text] to be mutually comparable.)

We must point out, however, that although necessary for concluding such shape invariance, this finding is not, of course, sufficient. Later, we suggest an alternative model that explicitly incorporates dependence of iconic-decay shape on stimulus duration, and we show that this model's prediction is essentially indistinguishable from horizontal parallelism.

Equation 11's fitting correctly is consistent with a strong version of the relatively weak model that we presented earlier in this article. Recall that the only constraints placed on the function $h(I)$ were that it be nonnegative, that it be equal to 1.0 when $I(t) = 0$, and that it approach 0 as $I(t)$ approaches 1.0. Similarly, the only constraint placed on the function $m(I)$ was that it be monotonically increasing. Equation 11, however, is a consequence of stronger forms of $h(I)$ and $m(I)$: that $h(I) = 1.0 - I$ and that $m(I) = I$.

Iconic Decay Is Not Generally Exponential

We noted at the onset of this article that iconic decay has often been assumed, either implicitly or explicitly, to be exponential. We have shown, however, that the exponential-decay hypothesis fails for at least 6 of the 8 observers in our experiment. We note that in addition to being interesting in and of itself, rejection of exponential decay is methodologically significant, as it indicates that our data are sufficiently powerful to reject a reasonable model.

GENERAL DISCUSSION

Let us summarize the general conclusions that we can make thus far.

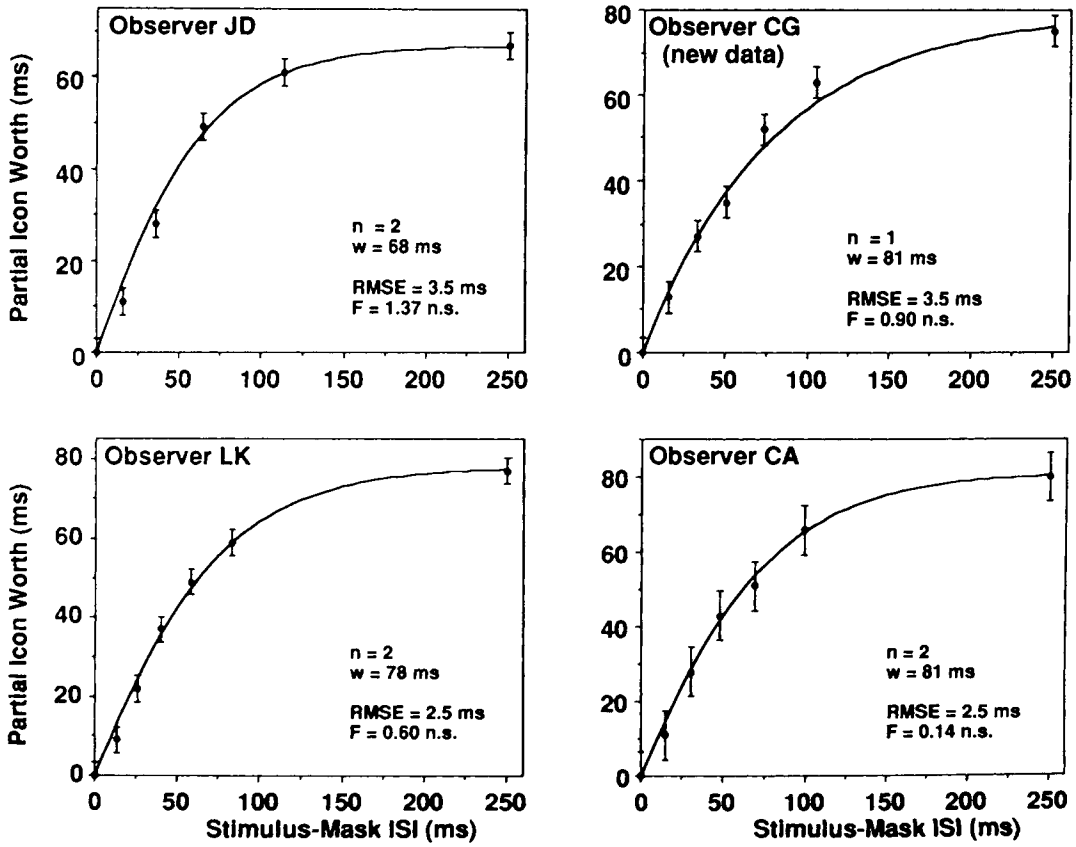

Figure 9. Best fitting gamma functions for female observers (JD, CG, LK, and CA).

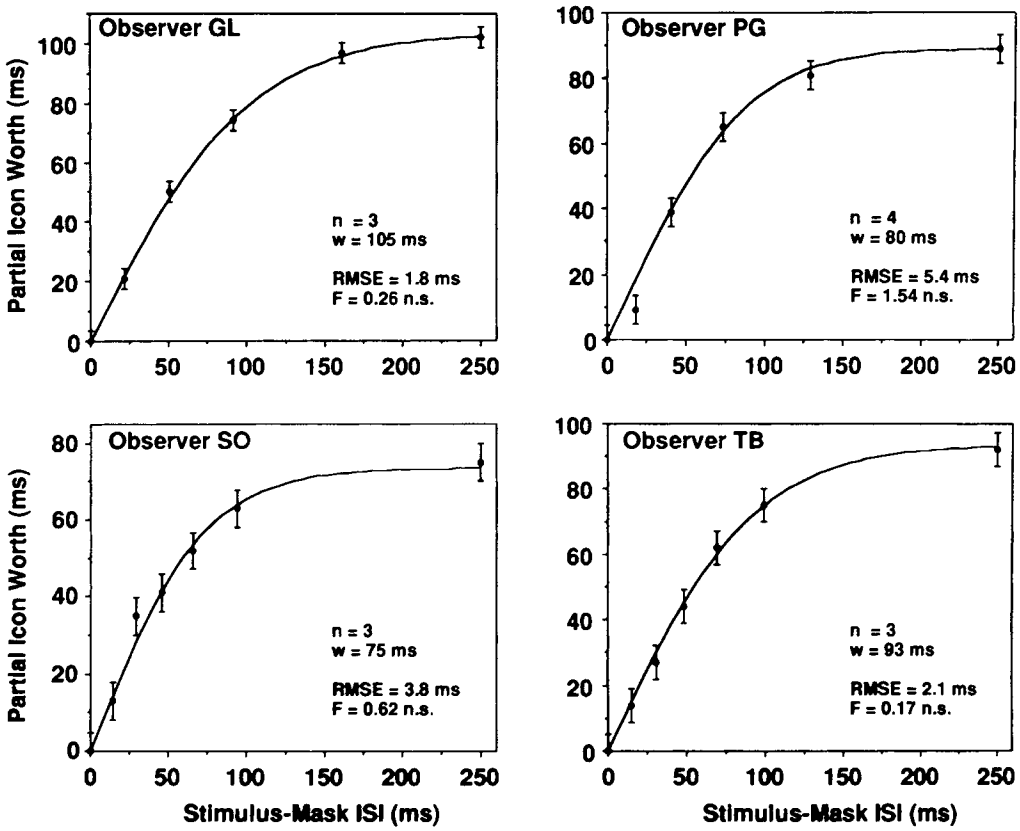

Figure 10. Best fitting gamma functions for male observers (GL, PG, SO, and TB).

First, in Experiment 1, we replicated the Loftus et al. (1985) data with a number of improvements. The data have more statistical power; we have assessed the null hypothesis of horizontally parallel immediate- and delayed-mask performance curves through standard statistical procedures, and we determined icon worths for individual observers. Second, in Experiment 2, we found families of horizontally parallel performance curves for a range of stimulus-mask ISI values. This finding confirms the prediction embodied in Equations 10 and 11 and is consistent with the proposition that the iconic-decay function's shape is invariant over stimulus duration. Third, we have rejected the hypothesis that iconic decay is exponential for 6 of 8 observers.

In the remainder of this section, we address two issues. First, we suggest a more general candidate decay function, gamma decay, that accounts for the present data. Then we consider more general models that are designed to account for the entire time course of the kind of $a(t)$ function shown in Figure 1.

Gamma Decay

A generalization of exponential decay is *gamma decay.* Gamma decay incorporates the notion of a series of stages: The output of each stage constitutes the input to the next one, and decay from each stage is exponential (see McClelland, 1979; McGill, 1963; McGill & Gibbon, 1965; Sperling, 1964; Watson, 1986). Decay from the nth stage, which we term $b(t - d)$, is $b(t - d) = e^{-(t-d)/\tau}\sum[(t/\tau)^i/i!]$ where τ is a scaling parameter, and the sum is from $i = 0$ to $i = (n - 1)$; thus n is a parameter restricted to integer values of 1 or greater.[9, 10] Note that when $n = 1$, gamma decay reduces to exponential decay.

Fit of Gamma Decay

For each of the 8 observers, we found the best fitting values of τ and n. The data (partial-worth curves) for each of the observers, along with the best fitting gamma-decay predictions, are shown in Figure 9 and 10. The best fitting parameters are expressed as n and w. The latter, which is equal to $n\tau$, is an estimate of the total area under $b(t - d)$, and as before is interpreted as the icon's total worth.

[9] Actually, this is one version of a gamma distribution called an *Erlang distribution.* In an Erlang distribution, the parameter n is restricted to integer values.

[10] There is a small terminological difficulty here. The equation that we have described is technically a gamma survival function, which is one minus a cumulative gamma distribution. The term *gamma function* is reserved for a particular mathematical function that appears in a variety of places. We use the terms *gamma decay function* and gamma function informally to refer to a gamma survival function.

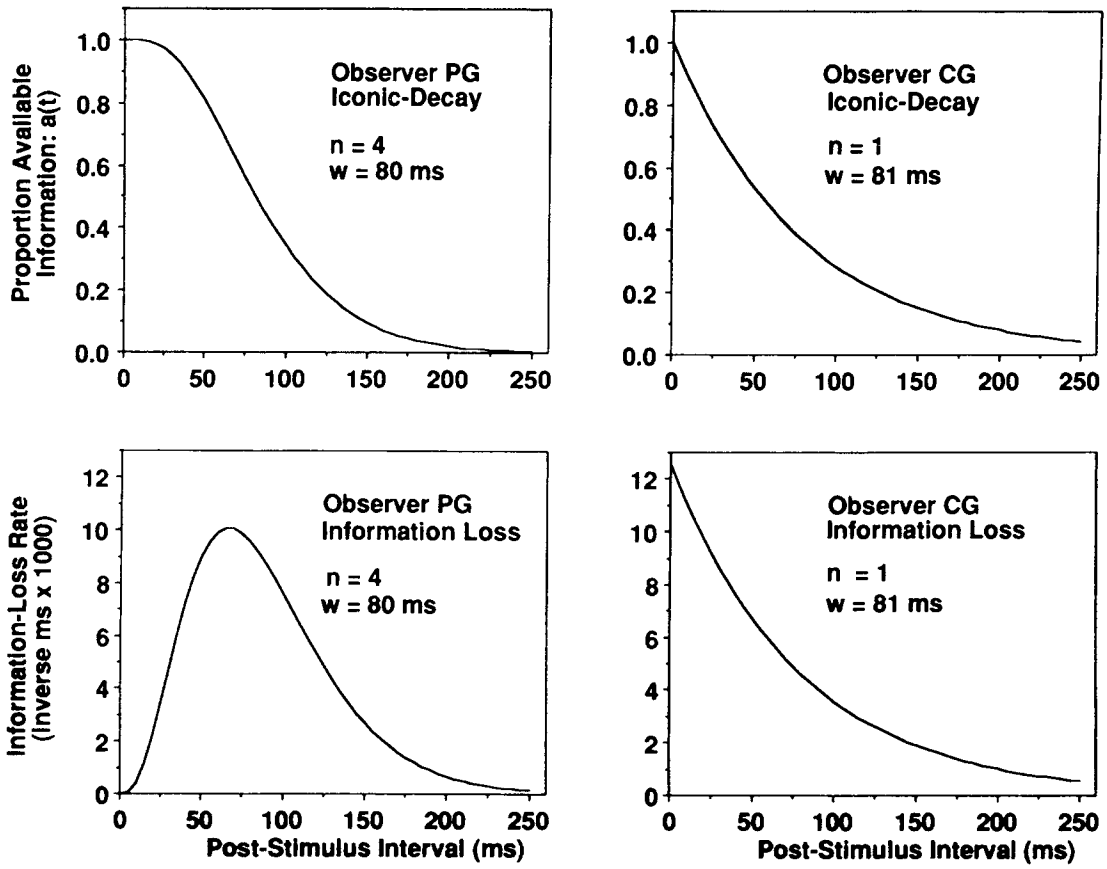

Figure 11. Top panels: Iconic-decay ($b[t - d]$) functions for Observers CG and PG. Bottom panels: Derivative of the iconic-decay functions, which are interpreted as information-loss-rate functions.

All of the fits are statistically nonsignificant. The best fitting n values range from 1.0 for CG to 4.0 for PG. All four n values for the female observers are either 1 or 2, whereas all four n values for the male observers are either 3 or 4. Figure 11 shows additional fit information for the two extreme observers: CG and PG. For both, Figure 11 shows the best fitting decay function ($b[t - d]$, the derivative of the best fitting $B[t - d]$) along with the *derivative* of the decay function, which represents information *loss rate* as a function of time since stimulus offset. Modal information loss rate occurs immediately for CG (whose decay is exponential), whereas modal information loss rate is delayed by about 60 ms following stimulus offset for PG.

Individual Differences

If taken literally, the implications of the gamma-decay fits are odd. Recall that within a gamma-decay model, the parameter n represents the number of exponential-decay stages that a representation undergoes prior to the available-information stage. Yet, the estimates of n vary from 1 to 4 across the 8 observers. It seems curious that such a presumably fundamental set of cognitive components would be built differently in different people.

Weichselgartner and Sperling (1985) reached similar conclusions about individual differences. They measured the apparent brightness of a briefly presented (31-ms) square-wave grating by having observers match stimulus brightness to an adjustable-luminance reference stimulus at various times before, during, and after stimulus onset. They found that the decay portions of what they referred to as the *temporal-brightness-response* (TBR) functions were quite variable across 3 observers and concluded that "these very significant individual differences are not adequately characterized by a single duration parameter, such as the decay constant of an exponential decay function. In fact, the TBRs do not seem to derive from any generic function. To describe visible persistence requires specification of a function, the TBR" (p. 720).

Constraining the Value of n Across Observers

There is a possible resolution to the puzzle of different observers apparently having different numbers of exponential-decay stages. We have noted that the *best fitting* values of n vary from 1 to 4 across the 4 observers. It is possible, however, that these differences are due to statistical variation and that all observers' data could be adequately accounted for with a single n value.[11]

[11] We thank Dave Irwin for this suggestion.

We have already seen that this cannot be accomplished with $n = 1$ (simple exponential decay). We found the best fitting gamma-decay functions constraining n to be constant across observers for n values of 2–5. The results are shown in Table 5, which provides RMSEs and F values for all observers for each of these n values.

For no n value are the fits nonsignificant for all observers. However, for $n = 2$, the only significant deviation is one of CG's two data sets. For $n = 3$, both of CG's data sets deviate significantly. For $n = 4$ and $n = 5$, additional data sets deviated significantly. Accordingly, the conclusions are ambiguous. It is possible that gamma decay with $n = 2$ exponentially decaying stages fits adequately and that CG's one significant deviation is a statistical aberration.

An additional view of the $n = 2$ gamma-decay fit is provided in Figure 12. Figure 12, which is analogous to Figure 8, shows normalized individual observer data fit to gamma decay with $n = 2$. It is evident that (unlike the case of exponential decay) there is no systematic deviation between data and model. Moreover, the overall fit is statistically nonsignificant, $F < 1$.

Accounting for the Entire $a(t)$ Function

As noted earlier, Weichselgartner and Sperling (1985) found individual differences in the decay portion of the TBR functions. These researchers also found substantial *rise-time* differences. Available-information rise time, depicted in the left portion of our Figure 1 $a(t)$ function, has been given short shrift in our (and the literature's) accounts of available information; iconic decay has customarily stolen the focus when available-information functions are considered. In this section, we discuss the shortcomings of the Figure 1 $a(t)$ representation, and we then consider an alternative model that provides a unified account of the entire $a(t)$ function.

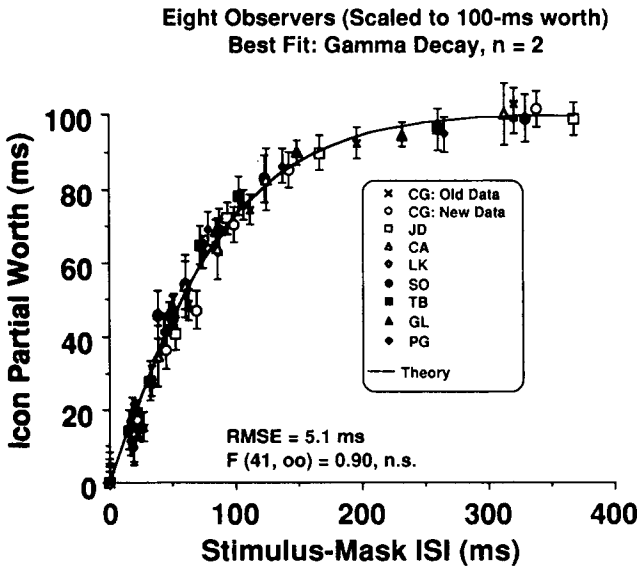

Figure 12. Best fitting gamma-decay function (constrained to $n = 2$) for 8 observers. (Observers' data have been scaled [as described in the text] to be mutually comparable.)

A Preliminary Remark

Before discussing our $a(t)$ representation, we wish to make a remark in defense of the general model that we proposed earlier. The model can be construed as having two components: first, the original available-information $a(t)$ function and second, everything else—the information-acquisition function, stored information, and the relation of stored information to memory performance. The two parts are modular in that they do not depend on each other; in particular, the post-$a(t)$ processes simply accept $a(t)$, whatever it is, and operate on it. Accordingly, alternative models can be proposed for generating $a(t)$ without disturbing the rest of the model.

Arbitrary a(t) Functions

Description of at least part of the $a(t)$ function—the iconic-decay part—has been a major purpose of the research described in this article. Although we have succeeded in deriving an empirically valid description of this function (see top panels of Figure 11), we have arrived at it in kind of a bootstrapping manner. We began with a made-up form of $a(t)$, depicted in Figure 1, that seemed reasonable on the basis of common sense, past data, and conventional wisdom. The critical features of this form were that (1) all information becomes available to the subsequent information-acquisition process fairly quickly following stimulus onset, and (2) there is only a single iconic-decay function that is tacked onto the end of the stimulus-present function, independent of stimulus duration (see Figure 4). Given this starting point, our Experiment 2 data allowed us to estimate the shape of the curve's iconic-decay portion.

However reasonable it may seem, though, and however good the data fit that emerges from it, the function depicted in Figure 1 is unsatisfyingly arbitrary in several ways. First, the genesis of the rising (left) portion of the curve is entirely unspecified. Second, the rising portion of the curve is in no way related to the decaying (right) portion of the curve. Third, no part of the curve issues from any more fundamental principles. It is a consequence of these deficiencies that our candidate iconic-decay functions were somewhat arbitrary.

A Linear-Systems Approach

There is a different way of arriving at an $a(t)$ function that is less arbitrary. This approach, which is based on linear-systems theory, is used to account for relatively low-level visual phenomena, such as Bloch's law and critical flicker frequency (for detailed accounts, see Sperling, 1964; Watson, 1986).

The Impulse-Response Function and Linearity

The linear-systems approach, whose goal is to describe the *system response* (e.g., Figure 1b) to some arbitrary *input function* (Figure 1a), is embodied in two principal assumptions. First, an instantaneous flash of light (an "impulse") is assumed to cause a nervous-system response, some facet of which is available to the subsequent processes that are respon-

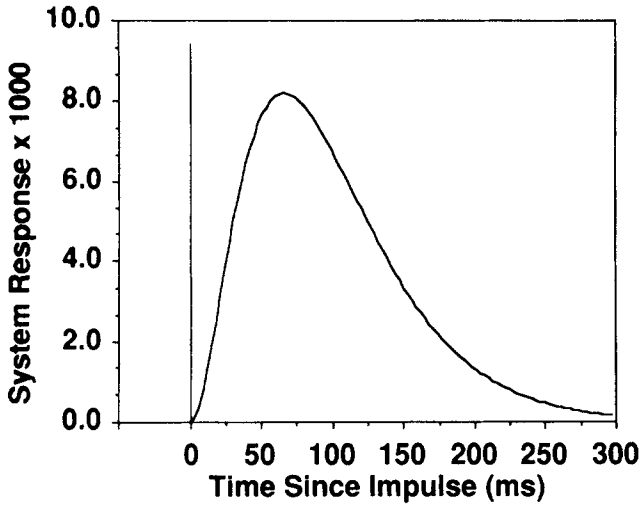

Figure 13. A gamma impulse-response function. (Vertical line at left represents the impulse, which occurs at time $t = 0$.)

sible for making a detection or a judgment. The system response is temporally blurred, as shown in Figure 13. The form of this *impulse-response function* is not entirely agreed on, but it is often assumed to be a gamma function (which is what is shown in Figure 13).

The second assumption is that the system is *linear.* This means that any input function can be decomposed into a series of scaled impulses and that the system response is the sum of correspondingly scaled impulse-response functions. More concisely, the system response is the convolution of the input function with the impulse-response function.

Letting $a(t)$ be a linear-system response function constitutes a simple way of generating the entire range of $a(t)$ from simple, plausible, and well-understood principles. Will it account for our data?

Application to the Present Paradigm

In the present experiments, the input function is very simple; it is the step function shown in Figure 1a. The response function is correspondingly simple:

$$a(t) \begin{matrix} = R(t) & \text{for } t \leq d \\ = R(t) - R(t-d) & \text{for } t > d, \end{matrix} \qquad (14)$$

where $R(t)$ is the integral of the impulse-response function.

Figure 14 shows the $a(t)$ functions that issue from Equation 14 for six duration values that are typical of those used in the present experiments. We used a gamma impulse-response function with $n = 3$ and $\tau = 33$. It is apparent that the resulting $a(t)$ curves differ in two ways from what we have been assuming (cf. Figures 1 and 4). First, each Figure 14 curve's rise time is approximately equal to its decay time. This is a fundamental characteristic of the system; it can be seen in Equation 14, in which the rise, embodied in $R(t)$, and decay, embodied in $R(t-d)$, are mirror images. Second, the

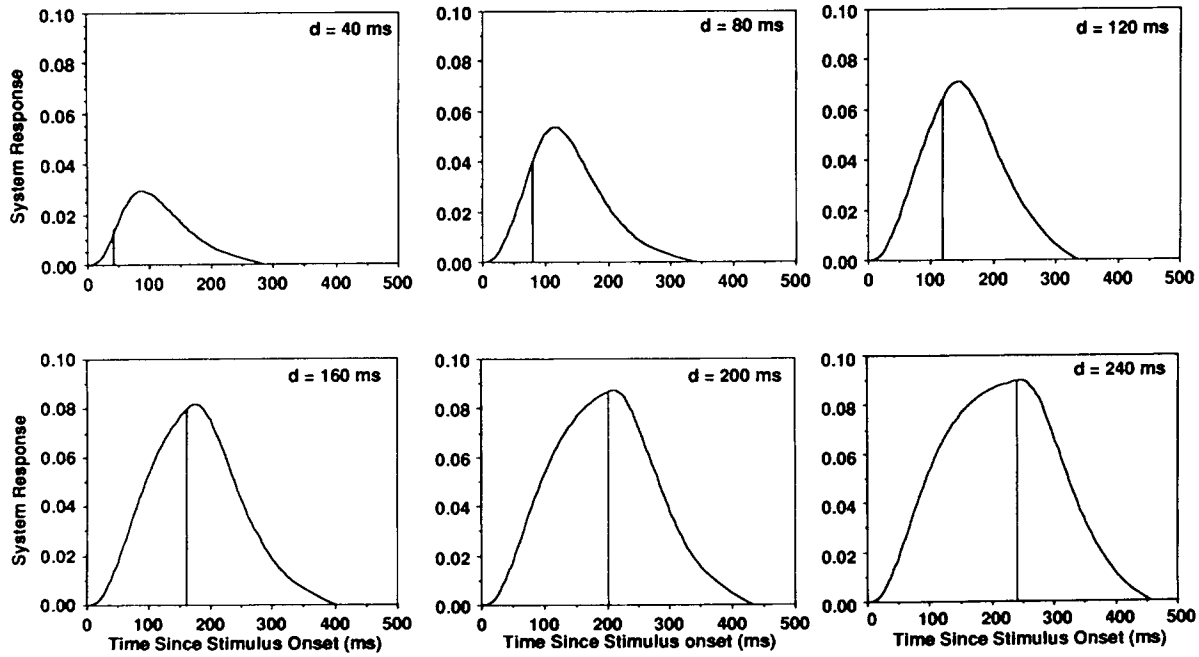

Figure 14. System responses to six step-input functions varying in duration from $d = 40$ to $d = 240$ ms. (Vertical line for each function denotes stimulus-offset time. Note that the iconic "decay" functions are not independent of stimulus duration.)

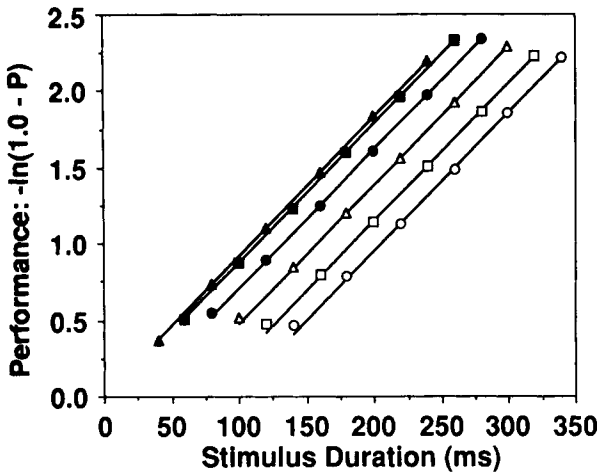

Figure 15. Prediction of the linear-systems model for the Experiment 2 data (compare with the real data in Figure 6).

shape of the iconic-decay portion of the curve is highly dependent on stimulus duration.[12]

Figure 15 shows the predicted data points generated by this model for the Experiment 2 paradigm along with the best linear fits. Superficially, at least, the linear model appears to do a very good job accounting for the data (cf. Figure 15 with Figure 6, which shows real Experiment 2 data for Observer JD).

However, this good fit is illusory. The problem can be unveiled by a careful comparison of Figures 6 and 15. It is apparent in Figure 6 that JD's lift-off (*d* intercept) value for her 250-ms ISI curve is about 65 ms. It is similarly apparent in Figure 15 that the model's corresponding curve passes through the origin. This latter characteristic—a zero-ms lift-off for a long-ISI performance curve—is a fundamental prediction of the linear-systems model that does not change with different parameter values. Accordingly, it is impossible for this simple version of the model to account for the present data.

Linear-Systems Variations

Other results from our laboratory more directly rule out a simple linear-systems account of data from the present paradigm. To illustrate, the linear-systems model predicts performance following a single, unmasked, 100-ms stimulus presentation to be the same as performance following two identical 50-ms presentations separated by a 100-ms ISI. In as-yet-unpublished data, we have disconfirmed this prediction (and other analogous predictions); performance following the single 100-ms presentation is superior.

Nonetheless, the Figure 15 prediction is tantalizingly close to being correct, and a linear-systems account of data from the present paradigm would have the advantages of (1) being tied to a model that is commonly used to account for low-level visual phenomena and (2) generating the $a(t)$ function from more fundamental principles. Accordingly, it seems fruitful to seek variations of the model that might work. One potentially profitable class of modifications incorporates the presumption that information acquisition does not begin until the system response has reached some criterion value, or until some criterion time has elapsed. We are currently evaluating these possibilities.

[12] One remark is in order here. Both differences we have just noted can almost be made to vanish by increasing the value of parameter n in the gamma function. The decay portion of the curve described by Equation 14, however, is essentially identical to the gamma functions that we fit to our data. Accordingly, the parameter values are constrained to those listed in Figure 11; a value of $n = 9$, for example, would produce an unacceptable fit of the decay portion of the curve to the data.

References

Adelson, E. H., & Jonides, J. (1980). The psychophysics of visual storage. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 486–493.

Averbach, E., & Coriell, H. S. (1961). Short-term memory in vision. *Bell Systems Technical Journal, 40,* 309–328.

Averbach, E., & Sperling, G. (1961). Short-term storage in vision. In C. Cherry (Ed.), *Symposium on information theory* (pp. 196–211). London: Butterworth.

Bogartz, R. S. (1990). Evaluating forgetting curves psychologically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 138–148.

Bowen, R. W., Pola, J., & Matin, L. (1974). Visual persistence: Effects of flash luminance, duration and energy. *Vision Research, 14,* 295–303.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183–228.

Di Lollo, V. (1984). On the relationship between stimulus intensity and duration of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 144–151.

Di Lollo, V., & Dixon, P. (1988). Two forms of persistence in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 671–681.

Efron, R. (1970a). Effect of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics, 8,* 231–234.

Efron, R. (1970b). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia, 8,* 37–55.

Eriksen, C. W. (1980). The use of a visual mask may seriously confound your experiment. *Perception & Psychophysics, 28,* 89–92.

Haber, R. N., & Standing, L. (1970). Direct estimates of the apparent duration of a flash. *Canadian Journal of Psychology, 24,* 216–229.

Hawkins, H. L., & Shulman, G. L. (1979). Two definitions of persistence in visual perception. *Perception & Psychophysics, 25,* 348–350.

Irwin, D. E., & Yeomans, J. M. (1986). Sensory registration and informational persistence. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 343–360.

Kowler, E., & Sperling, G. (1980). Transient stimulation does not aid visual search: Implications for the role of saccades. *Perception & Psychophysics, 27,* 1–10.

Loftus, G. R. (1985). Picture perception: Effects of luminance level on available information and information-extraction rate. *Journal of Experimental Psychology: General, 114,* 342–356.

Loftus, G. R., & Hanna, A. M. (1989). The phenomenology of spatial

integration: Data and models. *Cognitive Psychology, 21,* 363–397.
Loftus, G. R., Hanna, A., & Lester, L. (1988). Conceptual masking: How one picture steals attention from another picture. *Cognitive Psychology, 20,* 237–282.
Loftus, G. R., & Hogden, J. (1988). Picture perception: Information extraction and phenomenological persistence. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 139–191). New York: Academic Press.
Loftus, G. R., Johnson, C. A., & Shimamura, A. P. (1985). How much is an icon worth? *Journal of Experimental Psychology: Human Perception and Performance, 11,* 1–13.
Loftus, G. R., Kaufman, L., Nishimoto, T., & Ruthruff, E. (in press). Why is it annoying to look at slides with the room lights turned off? Effects of visual stimulus degradation on perceptual processing and long-term memory. In K. Rayner (Ed.), *Eye movements and cognitive processes.* New York: Springer-Verlag.
Loftus, G. R., Truax, P. E., & Nelson, W. W. (1986). Age-related differences in visual information processing: Quantitative or qualitative? In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structures over the life course* (pp. 60–78). Norwood, NJ: Ablex.
Mandler, J. M., & Parker, R. E. (1976). Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 38–48.
McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–324.
McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1). New York: Wiley.
McGill, W. J., & Gibbon, J. (1965). The general gamma distributions and reaction times. *Journal of Mathematical Psychology, 2,* 1–18.
Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.
Murdock, B. B., Jr., & Cook, C. D. (1960). On fitting the exponential. *Psychological Reports, 6,* 63–69.
Rumelhart, D. E. (1969). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology, 7,* 191–218.
Shibuya, H., & Bundsen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 591–600.
Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74,* 1–29.
Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5,* 19–31.
Sperling, G. (1964). Linear theory and the psychophysics of flicker. *Documenta Ophthalmologica, 18,* 3–15.
Sperling, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychologica, 27,* 285–292.
Stoddard, P. K., & Loftus, G. R. (1988). An IBM XT-compatible, computer-based, slide-projector laboratory. *Behavioral Research Methods, Instrumentation, and Computers, 20,* 541–551.
Townsend, J. T. (1981). Some characteristics of visual-whole report behavior. *Acta Psychologica, 47,* 149–173.
Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80,* 1–52.
Van der Heijden, A. H. C. (1978). *Short-term visual information forgetting.* Unpublished Ph.D. dissertation, University of Leiden, The Netherlands.
Watson, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 6-1–6-43). New York: Wiley.
Weichselgartner, E., & Sperling, G. (1985). Continuous measurement of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 711–725.
Wickelgren, W. (1977). Speed-accuracy tradeoff and information-processing dynamics. *Acta Psychologica, 41,* 67–85.
Yeomans, J. M., & Irwin, D. E. (1985). Stimulus duration and partial report performance. *Perception & Psychophysics, 37,* 163–169.

Appendix A

Proof That If $h(I) = (1.0 - I)$ and $M(I) = I$, Then Equation 3 Follows

Suppose that the function $h(I)$ is a linear function of I; $h(I) = (1.0 - I)$, and thus $r(t) = dI/dt = [a(t)(1.0 - I)]/c$, or

$$dI/(1.0 - I) = a(t)dt/c. \quad \text{(A1)}$$

Integrating both sides of Equation (A1), $-\ln(1.0 - I) = k_1 + A(t)/c$, where k_1 is a constant of integration and $A(t)$ is the integral of $a(t)$. Because m is the identity function, $P = I$. Thus,

$$1.0 - P = e^{k_1 + [A(t)/c]}. \quad \text{(A2)}$$

Note that $A(t)$ can be decomposed into three areas (see Figure 1): the area to the left of time $t = M$ (which we term k_2), the area to the right of d (which for a given stimulus-mask ISI is a constant that we term k_3), and the area in the middle (which is $d - M$). Making the substitution, $L = -(ck_1 + k_2 + k_3 - M)$, it follows that $1.0 - P = e^{-(d-L)/c}$, or $P = 1.0 - e^{-(d-L)/c}$. This completes the proof.

Appendix B

Derivations of the Equations for Acquired Information

Rate is the derivative of information with respect to time. Therefore, by Equation 5, $r(t) = dI/dt = [a(t)h(I)]/c$, or $c\,dI/h(I) = a(t)dt$. Integrating both sides,

$$H(I) = k_1 + A(t), \quad \text{(B1)}$$

where $H(I)$ is the integral of $c/h(I)$, $A(t)$ is the integral of $a(t)$, and k_1 is a constant of integration. Note that because $h(I)$ is greater than zero, $c/h(I)$, is both finite and also greater than zero. This in turn means that the integral of $1/h(I)$, $H(I)$, is monotonically increasing and, accordingly, has an inverse, $H^{-1}(I)$ (which, of course, is also

monotonic). Applying this inverse to both sides of Equation (B1),

$$I = H^{-1}[k_1 + A(t)]. \quad \text{(B2)}$$

Note that $A(t)$ can be decomposed into three parts: the area to the left of $t = M$ (which we term k_2), the area to the right of d, which is $B(t - d)$, the integral of the iconic-decay function, $b(t - d)$, and the area between M and d, which is $(d - M)$. Thus, Equation B2 may be rewritten as

$$I = H^{-1}[k_1 + k_2 + d - M + B(t - d)]. \quad \text{(B3)}$$

Letting $k = k_1 + k_2 - M$, Equation B3 may be written as

$$I = H^{-1}[k + d + B(t - d)]. \quad \text{(B4)}$$

In the case where $t = d$, $B(t - d) = 0$, and

$$I = H^{-1}[k + d]. \quad \text{(B5)}$$

Equations B3 and B5 thus constitute the model's equations for acquired information.

Appendix C

Proof That Horizontally Parallel Performance Curves Imply Equal c Values

Consider Equation 3. Suppose that any two performance curves (1 and 2) have equal c values but different L values. Thus,

$$\text{Curve 1: } P_1 = 1.0 - e^{-c(d - L_1)} \quad \text{(C1)}$$

and

$$\text{Curve 2: } P_2 = 1.0 - e^{-c(d - L_2)}. \quad \text{(C2)}$$

Determine the durations, d_1 and d_2, at which P_1 and P_2 are equal by setting Equations (C1) and (C2) to be equal: $d_1 - L_1 = d_2 - L_2$ and $d_1 - d_2 = L_1 - L_2$. The duration difference $(d_1 - d_2)$ that corresponds to equal performance is a constant $(L_1 - L_2)$, which means the curves are horizontally parallel. Thus, equal c values imply horizontally parallel curves.

Now drop the constraint that the c values be equal. Thus,

$$\text{Curve 1: } P_1 = 1.0 - e^{-c_1(d - L_1)} \quad \text{(C3)}$$

and

$$\text{Curve 2: } P_2 = 1.0 - e^{-c_2(d - L_2)}. \quad \text{(C4)}$$

From Equations C3 and C4, determine the relation between durations d_1 and d_2 that produces equal P_1 and P_2 values: $c_1(d_1 - L_1) = c_2(d_2 - L_2)$, or

$$d_2 = L_2 + [c_1(d_1 - L_1)/c_2]. \quad \text{(C5)}$$

The derivatives of P_1 and P_2 with respect to d are

$$dP_1/dd = c_1 e^{-c_1(d_1 - L_1)} \quad \text{(C6)}$$

and

$$dP_2/dd = c_2 e^{-c_2(d_2 - L_2)}. \quad \text{(C7)}$$

Parallel curves imply equal values of the derivatives when performance is equal. Accordingly, we let Equations C6 and C7 be equal when the relation expressed in Equation C5 holds. Thus,

$$c_1 e^{-c_1(d_1 - L_1)} = c_2 e^{-c_2(d_2 - L_2)}$$

or

$$c_1 e^{-c_1(d_1 - L_1)} = c_2 e^{-c_2[L_2 + [c_1(d_1 - L_1)/c_2] - L_2]},$$

which can be reduced to

$$c_1 e^{-c_1(d_1 - L_1)} = c_2 e^{-c_2 L_2 - c_1(d_1 - L_1) + c_2 L_2}$$

or

$$c_1 e^{-c_1(d_1 - L_1)} = c_2 e^{-c_1(d_1 - L_1)},$$

which implies that $c_1 = c_2$. This completes the proof that horizontally parallel curves imply equal c values.

Received October 9, 1990
Revision received February 5, 1991
Accepted February 13, 1991 ■