Sensory and Cognitive Components of Visual Information Acquisition

Thomas A. Busey and Geoffrey R. Loftus

The authors describe a theory of visual information acquisition and visual memory. The theory has 2 major components. First, the visual system's initial *sensory response* to a short-duration, low-contrast stimulus is generated by a linear, low-pass temporal filter that operates on the stimulus's temporal waveform. Second, *information* is *acquired* from a stimulus through an independent-sampling process whose sampling rate at time t following stimulus onset is jointly proportional to (a) the magnitude by which the sensory response exceeds some threshold and (b) the proportion of still unacquired information. The theory was successfully tested in 5 variants of a digit recall task in which temporal waveform of the stimulus was systematically manipulated. In a final experiment, the theory simultaneously accounted for performance in detection and identification tasks. Implications for visual information processing, low-contrast detection, and binocular combination of information are discussed.

Consider an observer looking at a visual stimulus with the intent of being able to remember it later on. The stimulus might be small or large, dim or bright, as simple as a single digit or as complicated as a natural scene. Regardless of the stimulus's nature, however, there are some fundamental processes that must always occur. First, there must be some registration of the stimulus in the visual system. Second, this registration must produce some sensory response that is intimately tied to stimulus presence. Third, based on this sensory response, information about the stimulus must be acquired and stored in some more permanent memory where it can be used as a basis for further processing or the subsequent demonstration of knowledge about (remembering) the stimulus.

The major goal of this article is to present a simple theory of these fundamental processes. The article is organized as follows. First, we describe the task that we are trying to account for, along with past data that have been obtained in this task and our rationale for why the task is useful for studying sensory processes, information acquisition, and memory. Second, we propose a theory to account for data in this task. Third, we present a very simple experiment designed to demonstrate the generality of certain crucial aspects of these past data and to elucidate the meanings of the theoretical parameters. Fourth, we present four experiments designed to test specific aspects of the theory.

Thomas A. Busey and Geoffrey R. Loftus, Department of Psychology, University of Washington.

Correspondence concerning this article should be addressed to Thomas A. Busey, Department of Psychology, University of Washington, Seattle, Washington 98195. Electronic mail may be sent to Busey@u.washington.edu.

Finally, having demonstrated that the theory provides an adequate description of our task, we extend it to a new task, the detection of the presence of an alphanumeric character.

A Digit Recall Task

The task to which we apply our theory is simple: Four digits are presented to an observer whose job is to report as many of them as possible, in their correct order, guessing if necessary. The basic performance measure, p, is the proportion of correctly reported digits, in their correct locations, adjusted for the guessing probability of 0.10.[1]

Performance Curves

In these experiments, stimulus exposure duration was varied over a range of approximately 20–200 ms. It is useful to present the data in the form of a *performance curve,* which is a curve relating performance to exposure duration. Loftus, Duncan, and Gehrig (1992) reported that such curves can be described almost perfectly by the function,

$$p = \begin{cases} 0 & (d < L) \\ Y(1.0 - e^{-(d-L)/c_r}) & (d \geq L) \end{cases} \tag{1}$$

which is illustrated in the top panel of Figure 1. Here d is exposure duration, and Y, L, and c_r are free parameters: c_r is the exponential growth constant for the regression model, L (for *liftoff*) is the maximum duration that gives chance-level performance (i.e., the duration at which performance "lifts off" from chance), and Y is asymptotic performance. As much as possible we ignore Y in this article. The reason for this is that the four-digit stimuli in our task are easily within the span of short-term memory. Therefore, given sufficient duration, a subject could, in principle, be perfect in this task, that is, Y could be 1.0 (as it was in all of the Loftus et al., 1992, experiments). For this rea-

---

[1] The guessing formula is p = $(x - 0.1)/0.9$, where x is the raw proportion correct and p is the corrected proportion.
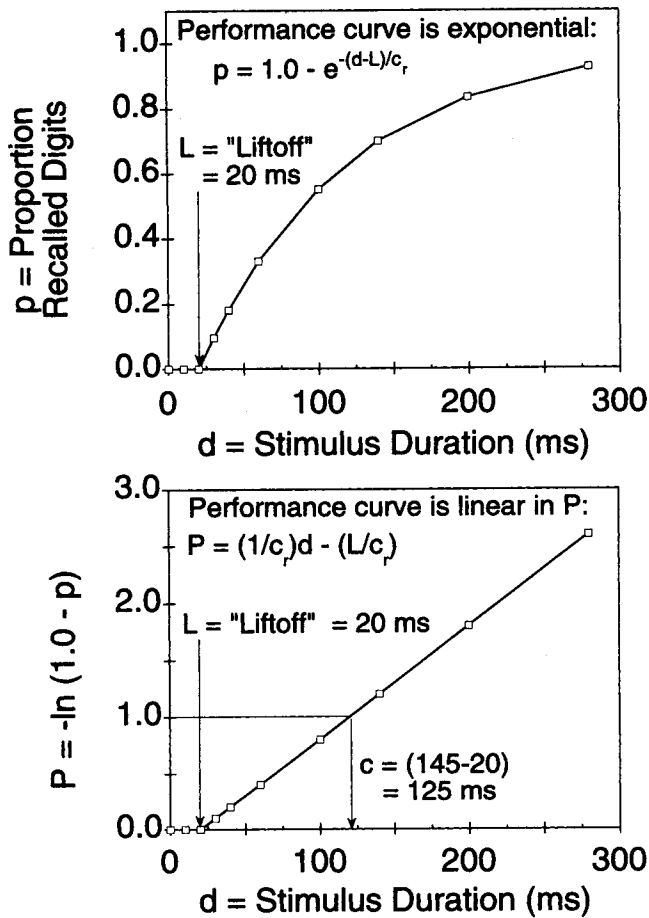
Figure 1. Top panel: a typical exponential performance curve for a many-duration hypothetical experiment. Bottom panel: The same data are transformed by $P = -\ln(1.0 - p)$, which produce linear functions. Linear regression techniques compute the best fitting line for the above-chance points and estimate the liftoff and c_r parameters.

son, a value of $Y < 1.0$ represents keypress errors, lack of vigilance, and soon, and is theoretically uninteresting from the perspective of our study. In all experiments, we estimated Y and removed its effects from the data by using Equation 1 (effectively setting Y to 1.0, as illustrated in Figure 1).

Given that performance curves can be adequately described by Equation 1, it is convenient to define a new dependent variable, P, as $-\ln(1.0 - p/Y)$. With P as the performance measure, Equation 1 can be rewritten as follows:

$$P = \begin{cases} 0 & (d < L) \\ d/c_r - d/L & (d \geq L). \end{cases} \quad (2)$$

Thus, in terms of P, post-liftoff performance is linear with duration with a slope of $1/c_r$ and a d-intercept of L ms, as illustrated in the bottom panel of Figure 1.

Appropriateness of the Digit Recall Task for Studying Perception and Memory

We argue that this digit recall task is useful for studying perception and the relation between perception and memory in

several ways. First, the task, although simple, still requires the basic components involved in complex cognitive processing, including sensory registration, acquisition and transfer of information to a more permanent memory, and generation of a response based on a memory representation (in contrast, for example, to a simple detection task). Second, the task does not require extraneous processes (such as registration of a probe in a partial report task); hence, a theory of the task can be focused and to the point. Third, the observed linearity of the performance curves (see Figure 1) strongly constrains the nature of the theories that can account for performance in the task. Finally, the parameters c_r and L have, as will become evident, direct and intuitive counterparts in the theory that we are presenting.

Variants of this task, along with variants of Equation 1 to describe them, are common in the information-processing literature. For example, Shibuya and Bundesen (1988), Townsend (1981), Rumelhart (1970), and Schurman, Eriksen, and Rohrbaugh (1968) all presented one or more characters for brief durations, usually followed by a visual noise mask. Often the data are characterized by models that make predictions similar to that embodied in Equation 1. Both Shibuya and Bundesen and Townsend explicitly included a model parameter, analogous to L, that represents the "processing delay." Rumelhart (1970) concluded that such a model parameter was probably necessary, although he did not explicitly include one. In past investigations, such a model parameter has been included simply to account for some aspect of the data; that is, it was not derived from more fundamental principles. Our first experiment (described in a later section) was motivated by the ubiquity of this processing delay and a desire to model it by using principles that are intimately tied to known visual-system behavior.

Theory

In this section we develop a theory that generates quantitative predictions for all aspects of the Figure 1 performance curve and for other similar paradigms. The theory conjoins two subtheories or models that have been used to describe low-level perceptual and higher level cognitive processes. We provide a brief overview of these models and then describe them in detail. The Glossary in Appendix A summarizes our notations.

Overview

The first model, termed the *linear-filter model,* generates an initial sensory representation in the nervous system. Briefly, this model represents the stimulus as a temporal waveform, that is, as a function relating stimulus contrast to time since stimulus onset. The model then assumes that the initial stages of the visual system act as a low-pass linear temporal filter on the stimulus waveform to produce a *sensory response function,* referred to as $a(t)$. The sensory response function relates the magnitude of some stimulus-signaling neural process to time since stimulus onset.

The second model, termed the *acquisition-rate model,* describes how stimulus information is acquired and transferred to some more permanent memory and used as a basis for a memory response. The acquisition-rate model has been described by

Loftus and his colleagues to account for temporal integration tasks (Loftus & Hanna, 1989), relatively high-level picture-processing tasks (Loftus & Hogden, 1988; Loftus, Hanna, & Lester, 1988), and for the digit recall tasks used in our experiments (Loftus et al., 1992; Loftus & Busey, 1992; Loftus, Busey, & Senders, 1993; see also Di Lollo & Dixon, 1992). The acquisition-rate model begins with the presupposition that the stimulus engenders *some* sensory response function that rises following stimulus onset and decays following stimulus offset.[2] It then assumes that (a) information is randomly acquired from the stimulus and placed in a memory store where it forms the basis of subsequent memory performance, and (b) subsequent memory performance is determined by the amount of acquired information.

We now describe each component in detail. We begin by showing how the sensory response function is generated by the linear-filter model. We then use the acquisition-rate model to derive predictions by combining the sensory response function with already acquired information.

A Linear-Filter Model to Generate a Sensory Response Function

The linear-filter model describes the visual system's initial response to a briefly presented stimulus.

The Stimulus Contrast Function: f(t)

The stimulus can be defined as a *stimulus contrast function*, $f(t)$, that relates physical stimulus contrast to time t since stimulus onset as indicated in Figure 2, top panel. The $f(t)$ shown in
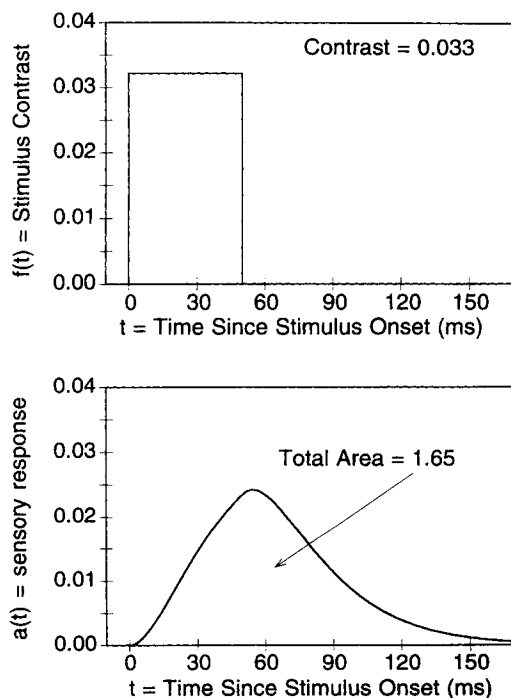

Figure 2. The $f(t)$ (top panel) and $a(t)$ (bottom panel) curves for a 50-ms presentation at contrast 0.033. The total area under the $a(t)$ curve = 50 ms × 0.033 = 1.65.

Figure 2 is rectangular, corresponding to the actual stimuli used in many experiments. However, the theory applies without modification or additional assumptions to any arbitrary non-negative $f(t)$.

Assumption 1: The Impulse Response Function

The stimulus contrast function, $f(t)$, may be mathematically expressed as a series of instantaneous[3] impulses, each of which generates what is called an *impulse response function* (scaled by stimulus contrast). Along with others, we assume the impulse response function to be gamma function of the form,

$$g(t) = \frac{(t/\tau)^{n-1}e^{-t/\tau}}{\tau(n-1)!}, \tag{3}$$

where n and τ are free parameters: n is a positive integer, and τ is a positive real number. This function can be interpreted as representing the output of an n-stage system where the input to Stage 1 is the stimulus, the input to each of Stages 2 through n is output of the previous stage, and the output of each stage decays exponentially with decay constant τ. It will turn out that the fit of the theory to our data is relatively unaffected by the value of n. We set n to 2 for all of the theoretical fits associated with Experiments 1–4 and set n to 9 for fits associated with Experiments 5 and 6.[4]

This impulse response function represents a subset of all possible impulse response functions and differs from similar functions in the literature in that it is *monophasic*, that is, it has a single relative maximum. Others have proposed biphasic (Sperling & Sondhi, 1968; Watson, 1986) and even triphasic (Roufs & Blommaert, 1981) impulse response functions. Adding additional relative maxima to the linear-filter model would require additional assumptions for negative-going sensory response functions.

Assumption 2: The Sensory Response Function: a(t)

We define the sensory response function, $a(t)$, as the *convolution* of $f(t)$, the stimulus input function, and $g(t)$, the impulse response function:

$$a(t) = f(t)*g(t). \tag{4}$$

This definition is based on fundamental principles issuing from linear systems theory (see Groner, Bischof, & Di Lollo, 1988; Loftus et al., 1992; Sperling, 1964, 1979; Watson, 1986, for discussions of linear system applications to perceptual processes). Essentially, it results from the proposition that each of the impulses comprising $f(t)$ generates an individual impulse response

---

[2] Thus, only a weak definition of $a(t)$, the sensory response function, is needed by the acquisition-rate model. The linear-filter model's job is to specify $a(t)$ precisely.

[3] An impulse is a mathematical fiction defined to be a stimulus of infinitesimal duration, infinite intensity, and unit area.

[4] An n of 1 provides an exponential impulse response function and does not adequately predict the data. Early experiments used an n of 2 to improve computation time during parameter estimation. Later experiments were fit with an n of 9 to conform with generally accepted values in the vision literature.

function and that these impulse response functions are then summed to produce $a(t)$.

One implication of Equation 4 and our choice of a monophasic impulse response function is that the total area under the sensory response function, which we call $A(\infty)$, equals the total area under the stimulus contrast function, which we call $F(\infty)$, that is, $A(\infty) = F(\infty)$. This property does not hold for multiphasic impulse response functions, which may give negative-going impulse response functions and imply $A(\infty) \leq F(\infty)$.

Implicit in this definition of the sensory response function is the notion of a linear transducer function, which means that the internal effect of the stimulus is directly proportional to the stimulus contrast, $f(t)$. Others working in low-contrast visual detection and discrimination have assumed that the sensory response function is proportional to the square of contrast (Carlson & Klopfenstein, 1985; Nachmias & Sansbury, 1974). This has been generalized to a power law of the function by Legge and others (Legge, 1980; Gottesman, Rubin, & Legge, 1981) of the form $\Delta C \propto C^N$, where ΔC is the increment contrast threshold, and C is contrast. However, a linear transducer function has been successfully applied by Sachs, Nachmias, and Robson (1971) and Graham (1977). A quadratic transducer function was attempted for the current theory, with distinctly poorer fits.

In Experiments 1–4 we used a rectangular-function display shown in the top panel of Figure 2: To display the stimulus, the projector shutter opened, essentially instantaneously, remained open for the duration of the stimulus, and then closed, again essentially instantaneously. For a d-ms rectangular function whose maximum contrast is ϕ, the resulting $f(t)$ function is

$$f(t) = \begin{cases} \phi & (0 \leq t \leq d) \\ 0 & \text{(elsewhere)} . \end{cases}$$

The convolution of the impulse response function (or indeed any function) $g(t)$ with such a rectangular $f(t)$ function is quite simple. The resulting $a(t)$ function is

$$a(t) = \begin{cases} \phi G(t) & (t \leq d) \\ \phi[G(t) - G(t - d)] & (t > d) \end{cases} \tag{5}$$

where $G(x)$ is the integral of $g(x)$ from 0 to x. The $a(t)$ function shown in Figure 2 was generated from Equation 7 with $\phi = 0.033$, $d = 50$ ms, $\tau = 20$ ms, and $n = 2$.

Assumption 3: The Threshold Assumption

We make one additional assumption that is quite standard in the application of linear-filter models to sensory data: We assume there to be some threshold sensory response, Θ, such that information acquisition does not occur unless $a(t) > \Theta$. Thus, an effective sensory response function may thus be defined as

$$a_\Theta(t) = \begin{cases} a(t) - \Theta & (a(t) > \Theta) \\ 0 & (a(t) \leq \Theta) \end{cases} \tag{6}$$

where Θ represents the sensory threshold in units of stimulus contrast.

Acquisition-Rate Model

The acquisition-rate model uses the sensory response function as input. It generates an instantaneous information-acquisition rate, along with predictions for our tasks.

Assumption 4: Information Is Acquired Independently

Stimulus information is acquired by the observer and placed into some short-term store. In particular, the stimulus is conceived to be a bundle of features that the observer samples over time. Sampling is assumed to occur *independently*, such that the probability of acquiring one feature is independent of acquiring another feature. Models based on this assumption have appeared in Townsend (1981), Shibuya and Bundesen (1988), Rumelhart (1970), and Massaro (1970).

At time t following stimulus onset, some proportion, $I(t)$, of the total information has been acquired. We assume that the instantaneous information-acquisition rate at time t, designated $r(t)$, is the product of two quantities: first, the magnitude by which the sensory response exceeds the sensory threshold, $a_\Theta(t)$, and second, some monotonically decreasing function of already acquired information.

The *particular* monotonically decreasing function of already acquired information is implied by the logic of independent sampling. The raw feature-sampling rate, designated as $1/c_s$, remains constant; however, the rate at which new (i.e., previously unsampled) features are sampled decreases over time as the number of old (i.e., already sampled and stored) features increases. In particular, as shown in Appendix B, independent sampling implies—ignoring for the moment, the contribution of $a_\Theta(t)$—that new information is sampled at a rate of $[1.0 - I(t)]/c_s$.

We note also that $r(t)$ is, by definition, the derivative over time of $I(t)$. Thus,

$$r(t) = \frac{dI}{dt} = a_\Theta(t)\left[\frac{1.0 - I(t)}{c_s}\right]. \tag{7}$$

Appendix B also shows that, with this rate function, the equation relating total acquired information, which we designate $I(\infty)$, to the above-threshold area under $a_\Theta(t)$, $A_\Theta(\infty)$, becomes

$$I(\infty) = 1.0 - e^{-A_\Theta(\infty)/c_s}. \tag{8}$$

Figure 3 summarizes three major components of the model: the stimulus input waveform, $f(t)$; the resulting sensory response function, $a(t)$; and the information-acquisition rate function, $r(t)$.

Assumption 5: Performance Is Equivalent to Acquired Information

To provide predictions for our tasks, we make the simplifying assumption that p, the proportion of correct digits, is equal to $I(\infty)$, the total proportion of acquired information. Thus,

$$p = 1.0 - e^{-A_\Theta(\infty)/c_s}.$$

Note that, in terms of $P = -\ln(1.0 - p)$,

$$P = A_\Theta(\infty)/c_s. \tag{9}$$

Stimulus Waveform (40 ms Rectangular)

f(t) (% contrast): 0.08, 0.06, 0.04, 0.02, 0.00

Time Since Stimulus Onset (ms): 0 50 100 150 200 250

The Physical Stimulus

↓

Sensory Response Function

Sensory Response: a(t): 0.04, 0.03, 0.02, 0.01, 0.00

Sensory Threshold

Time Since Stimulus Onset (ms): 0 50 100 150 200 250

The Visual System's Response to the Physical Stimulus

↓

Information-Extraction Rate: r(t) in 1/ms

$r(t) = dI(t)/dt = a(t)[1.0 - I(t)]/c_s$

Extraction Rate: r(t): 0.015, 0.010, 0.005, 0.000

$I(\infty)$

Time Since Stimulus Onset (ms): 0 50 100 150 200 250

Above-Threshold Sensory Response is Proportional to Performance Via the Information-Extraction Rate
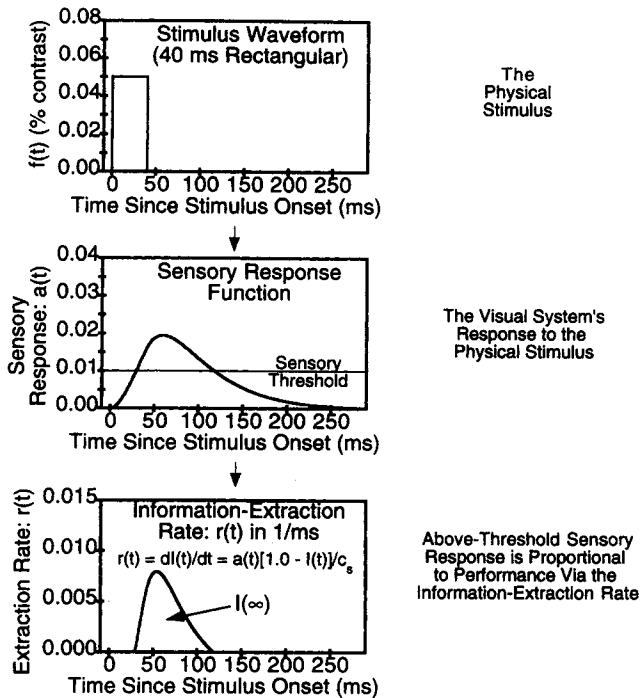
Figure 3. Summary of three model components. Top panel: $f(t)$, the stimulus contrast function. Middle panel: $a(t)$, the sensory response function generated by applying a linear temporal filter to $f(t)$. Bottom panel: $r(t)$, the information acquisition-rate function, proportional to the above-threshold magnitude of $a(t)$. c_s = extraction-rate parameter.

Equation 9 summarizes an important prediction: that performance is directly proportional to the above-threshold area under the $a(t)$ function.

Relationship to Probability Summation

The equation, $p = 1.0 - e^{-A_\Theta(\infty)/c_s}$, bears a striking resemblance to Watson's probability summation over time model (Watson, 1979). His model (using our notation) assumes that probability p is a function of the integral of the impulse response function raised to some power β:

$$p = 1 - (1 - \gamma)e^{-\int [a_\Theta(t)/c_s]^\beta dt},$$

where γ is a guessing parameter, and $a_\Theta(t)$ is the sensory response function. The current theory is similar to the probability summation model, with β set to 1, and the explicit assumption of a sensory threshold. The probability summation model implicitly assumes a threshold through the β parameter.

The probability summation model was applied to the current experiments, and produced qualitatively good fits. However, the current theory produced consistently better fits based on comparisons of root-mean-square errors (RMSEs).

Performance Curve Predictions

To illustrate the effect of the threshold, Θ, on the shape of the predicted performance curves, Figure 4 shows five $a(t)$ curves resulting from five different stimulus durations for contrast ϕ =

0.033, and the associated predicted performance curves. The top panel shows a truly linear system without a sensory threshold: here P, assumed proportional to $A_\Theta(\infty)$, is plotted as a function of d, and the curve is linear with zero intercept. The bottom panel shows the $a(t)$ curves that result with the sensory threshold Θ added, which violates the strict linearity of the system: here P, assumed proportional to $A_\Theta(\infty)$, is plotted as a function of d.[5] The curve is almost linear, with an above-zero d-intercept.

Generating Theoretical Predictions

To summarize, the theory has four free parameters: n and τ determine the impulse response function (although recall that n is set to either 2 in Experiments 1–4 or to 9 in Experiments 5 and 6); Θ is the sensory response threshold; and c_s is the proportionality constant that maps $A_\Theta(\infty)$ to P. To fit the theory to our data, we used grid search procedures to provide estimates of the three parameter values (τ, Θ, and c_s) and to assess the adequacy of the fits. The parameter values and RMSEs for all fits to all experiments are in Table 1. Theoretical predictions are shown as solid lines through the data points in all figures that depict experimental results. Additional procedures are required to apply the theory to each individual experiment, and these details are summarized in Appendix C.

Parameter Value Interpretations

The value of the sensory threshold parameter, Θ, was in the general range of 2% for Experiment 1 and for all experiments. From our theory's perspective, Θ is expressed in units of stimulus contrast. To see why this is so, recall that a stimulus is only recalled with an above-chance probability if $a(t)$ exceeds Θ. Now imagine a stimulus of constant contrast, ϕ, and indefinite duration. By Equation 5, as t increases, $G(t)$ would asymptote at 1.0, the total area under $g(t)$ and, correspondingly, $a(t)$ would asymptote at ϕ, the stimulus contrast. The threshold is thus interpreted as the lowest stimulus intensity such that $a(t)$ never exceeds threshold. An estimated value of around 2% implies that if a stimulus were shown at around 2% contrast level, performance would not rise above chance, no matter how long the stimulus was displayed.

Finally, as noted, c_s is the proportionality constant that maps $A_\Theta(\infty)$ to P. This c_s is related to the c_r of Equation 2, which provides the best linear fit between stimulus duration and P. The two differ because (a) duration and $A(\infty)$ differ by a factor equal to stimulus contrast, and (b) $A_\Theta(\infty)$ differs from $A(\infty)$ in that the former includes only above-threshold area.

Individual Data and Averaging Techniques

For Experiments 1–5 we focus, for expositional efficiency, on data averaged over observers. These data were obtained as follows. The means for each of the various conditions for each observer were averaged to provide the observed data points in the data figures. Each observer's data was then fit by the theory. These fits were averaged in the same manner as the observed
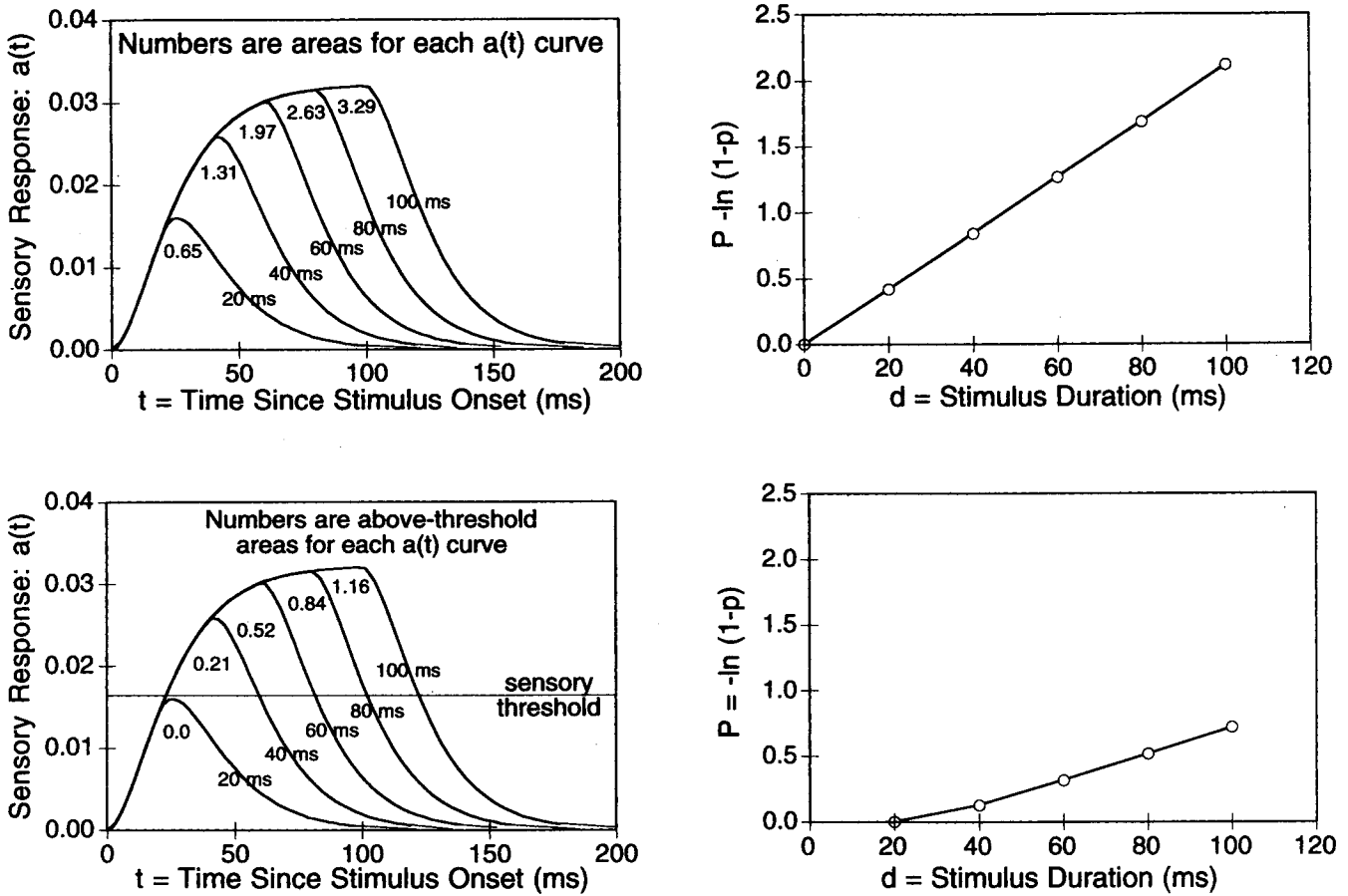
Figure 4. Upper left panel: five $a(t)$ curves resulting from 20-, 40-, 60-, 80-, and 100-ms presentations at contrast = 0.033. Upper right panel: predicted performance (P) for the five durations. Lower left panel: the five $a(t)$ curves with a sensory threshold. Lower right panel: predicted performance curve for the five presentations with a sensory threshold. The addition of the sensory threshold lowers overall performance and increases the liftoff duration from 0 to about 22 ms. p = proportion correct.

data and are shown as solid curves in the data figures. The RMSEs reported for the averaged-data figures were computed from the averaged observed data and the averaged model fits. To obtain average model parameters, the gridsearch procedures were applied directly to the averaged data.

In some cases there were qualitative or otherwise interesting individual differences. In these cases, we present data separately for individual observers.

Testing the Theory: Experiments 1–6

Each of Experiments 1–5 tests the theory as a whole, that is, each tests the conjunction of models that compose the theory. However, different experiments focus on different aspects of the theory: In particular, Experiment 1 partially replicates previous work; Experiment 2 focuses on the independent-sampling assumption; and Experiments 3–5 focus on the presumed sensory response threshold and the assumption that performance, P, is proportional to $A_\theta(\infty)$, the total above-threshold area under the $a(t)$ function. Experiments 1–5 fol-

low a similar theme: In all of them, the stimulus input function, $f(t)$, is manipulated.

Experiment 1: Does Performance-Curve Shape Depend on Whether a Mask Is Used?

Past work (e.g., Loftus et al., 1992, 1993; Shibuya & Bundesen, 1988; Sperling, 1963) has shown that performance curves in the digit-recall paradigm (a) were linear and (b) showed above-zero liftoffs. These two characteristics of the data are both fundamental predictions of the theory that we have described. Experiment 1 is designed to investigate the generality of these two characteristics. In particular, these past experiments all involved use of a random-noise mask to eliminate part or all of the iconic image that normally follows a briefly presented stimulus. Thus, performance has depended not only on seeing the stimulus but also on escaping the masking effect (e.g., Turvey, 1973; see Eriksen, 1980, for problems in data interpretation when a mask is used). Perhaps the characteristic performance-curve shape described

Table 1

Summary of Best Fitting Model Parameters: Experiments 1–5

Observer	τ	c_s	Detect c_s	Ident c_s	Θ	$\Theta(1)$	$\Theta(2)$	RMSE
				Experiment 1 ($n = 2$)				
AB	5.8	0.052			1.779			0.317
DA	9.1	1.008			2.527			0.089
KC	7.0	0.340			2.600			0.191
TB	10.9	1.727			1.264			0.090
Average	8.6	0.919			1.642			0.083
				Experiment 2 ($n = 2$)				
DA	8.8	0.549			2.704			0.215
KC	17.7	0.860			1.412			0.120
TB	46.7	1.542			0.564			0.067
AB	33.9	1.119			0.412			0.282
Average (DA, KC, TB)	10.7	0.594			2.050			0.086
				Experiment 2: Different first and second Θ				
AB	7.5	0.450				1.948	2.447	0.077
				Experiment 3 ($n = 2$)				
TB	11.6	0.774			2.598			0.042
TK	14.3	0.264			2.352			0.095
DA	14.2	0.599			2.135			0.069
Average	13.9	0.662			2.222			0.045
				Experiment 4 ($n = 2$)				
GW	24.1	1.796			1.638			0.029
TB	58.4	1.904			0.822			0.054
TK	16.6	0.865			1.990			0.046
Average	29.2	1.400			1.441			0.028
				Experiment 5 ($n = 9$)				
CW	6.5	1.610			3.879			0.0966
LM	7.3	1.222			4.108			0.1087
TB	6.4	1.809			4.024			0.0750
SB	6.4	2.633			5.783			0.0725
Average (CW, TB, LM)	6.6	1.523			4.037			0.0584
				Experiment 6 ($n = 9$)				
MB	8.3		0.229	1.674	3.256			0.0187
TB	4.3		0.396	1.681	3.200			0.0197
LM	8.1		0.437	0.614	4.361			0.0095
LM (1 c_s)	8.1		0.534	0.534	4.357			0.0140

Note. Mean gridsearches are derived from gridsearches on mean data. Average fits represent model parameters derived from fits on averaged data. Root-mean-square errors (RMSEs) are in units of $-\ln(1 - p)$ for Experiments 1–5 and in units of contrast for Experiment 6. Units are as follows: c_s and τ are expressed in milliseconds; Θs are expressed in percentages of contrast. In Experiment 2, $\Theta(1)$ was the first threshold, and $\Theta(2)$ was the second threshold. In Experiment 6, Detect c_s = detection task c_s; Ident c_s = identification task c_s; 1 c_s = a single c_s was used to model both tasks.

by Equations 1 and 2 depends on the mask's inclusion in the display. Our goal in Experiment 1 was to ascertain whether or not this was true.

Method

An experimental session consisted of a series of trials on each of which a four-digit array was presented. The observer's task on each trial was to report the digits in their correct locations, guessing if necessary.

Observers. Four observers participated in the experiment: the first author (TB), one undergraduate (KC), and two graduate students (AB and DA). All observers had participated in a minimum of 500 trials before beginning the experiment and were all familiar with the nature of the research.

Stimuli and apparatus. Stimuli were prepared as 35-mm slides each containing a 4 × 3 (Columns × Rows) array of black digits on a white background. Seventy-two such stimuli were prepared and used repeat-

edly. The $4 \times 3 \times 72 = 864$ digits consisting of all stimuli were selected randomly from the set of 10 digits, with the restrictions that (a) each digit appeared 96 times over the entire digit set and (b) no digit appeared more than twice in a given row. Each digit subtended a visual angle of 0.68° vertically and 0.35° horizontally. Digits were separated by 0.27° vertically and 0.76° horizontally. (The presentation view of Figure 10 shows a sample stimulus.) On a given experimental trial, either the top or the bottom four-digit row of one stimulus was the to-be-reported target. Target row was blocked over trials; accordingly, an observer always knew in advance which row was the target and therefore there was no positional uncertainty.[6]

Stimulus contrast was substantially reduced to avoid ceiling performance. Contrast reduction was accomplished by (a) attenuating stimulus luminance with a Wrattan neutral-density filter and (b) superimposing a uniform adapting field over the entire stimulus. A summary of luminances and contrasts for all experiments is provided in Table 2.

All stimuli in Experiments 1 and 2 were displayed by Kodak projectors equipped with Gerbrands tachistoscopic shutters. A random-access projector was used to display the stimuli, and standard carousel projectors were used to present the adapting field and a fixation slide that initiated each trial. All equipment was enclosed in a soundproof box. All stimulus display and response collection were under the control of an AT-compatible computer system described by Stoddard and Loftus (1988).

Design and procedure. Eight conditions were defined by eight stimulus durations, which ranged from 20 ms to 136 ms in 17-ms increments. Each observer participated in eight 72-trial blocks, thereby providing 72 observations per condition per observer. On each trial an

observer reported the four digits in either the top or bottom row. The to-be-reported row was constant within a block and alternated between blocks. Assignment of conditions to trials within a block and slide-presentation order also changed over blocks. Each block began with eight practice trials.

The sequence of events within a trial was as follows. First, there was a 500-ms fixation point, accompanied by a warning tone. Warning-tone frequency was 2,000 Hz or 500 Hz and reminded the observer which row (top 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 tone commanding the observer to respond. The observer then typed four digits into a response box, guessing when uncertain. Following the responses were four 150-ms feedback tones: each tone was 2,000 Hz if the corresponding digit had been correct and was 500 Hz if it had been incorrect. Following feedback was a 300-ms interval before the start of the next trial. The adapting field remained present throughout the experimental session.

Results and Discussion

The primary goal of Experiment 1 was to replicate the essential features of the Loftus et al. (1992, 1993) performance curves using unmasked stimulus presentations. Figure 5 shows the performance curve averaged over the four observers. Here the dashed line indicates the best linear fit,,22 and the solid line indicates the fit of our theory. The best linear fit was obtained by finding the values of the parameters c_r, L, and Y (from Equations 1 and 2) that produced the lowest RMSE between obtained and predicted data. Standard-error bars are shown for the Figure 5 data, as they are for all experimental data in this article; however, the error bars are often occluded by plot symbols.

It is clear that the linear fit to the Figure 5 data is quite good; the Pearson product–moment correlation is .988 for the mean data, and the worst individual correlation is .958. Although in this article we are primarily concerned with model predictions generated from our theory, we always provide the linear-fit parameters to performance curves, as these statistics are potentially useful for testing other theories besides ours. The linear parameter values along with the Pearson correlation values for performance curves from this and subsequent experiments are provided in Tables 3 through 6.

Despite the absence of the pattern mask, the performance-curve linearity remains, as does the above-zero d-intercept (lift-off). The predictions of our theory are shown as the solid line in Figure 5. The RMSE for the averaged data was 0.0833. Parameter values for all experiments for our theory can be found in Table 1. It is evident by inspection that our theory's fit is quite good.

Table 2
Luminances and Contrasts for All Experiments

Type of measurement	Digits	Fixation point
Experiments 1 and 2		
Background	19.730	31.070
Foreground	18.470	17.000
Contrast	0.033	0.293
Intensity	0.651	9.103
Experiments 3 and 4		
Background	7.272	7.272
Foreground	6.812	0.456
Contrast	0.037	0.883
Intensity	0.267	6.419
Experiments 5 and 6		
Background	7.272	7.272
Foreground	*	0.456

Experiment 5 contrasts by condition		
Maximum contrasts	Subjects TB, LM, and MB	Subject SB
Rectangular function	0.054	0.074
Ramp on	0.108	0.147
Ramp off	0.108	0.147

Note. All luminances are in candles/m². Contrast is defined as contrast $= (L_b - L_f)/(L_b + L_f)$, where L_b is the background luminance and L_f is the foreground luminance. Intensity is contrast * background luminance and has units of cd/m². Experiment 5 varies contrast by condition. Experiment 6 varies contrast to find a contrast threshold.

[6] A stimulus slide consisted of three rows rather than one row so that the slides could be used repeatedly, with different rows as the to-be-remembered array on successive viewings without the observers' being able to memorize the trial-to-trial sequences. In previous experiments using these stimuli, all three rows were used, with the relevant row changing on each block. However, the fixation point caused masking problems when the middle row was relevant; hence, in the present experiments we used only the top and bottom rows.
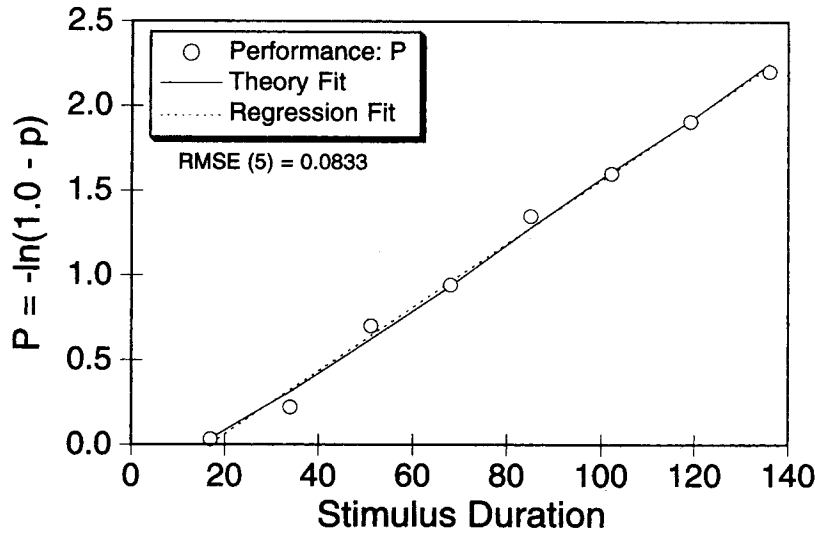
Figure 5. Experiment 1 data. The dashed line is the best fitting regression curve for the above-chance points. The solid curve is the best fit of the theory. As in all data figures in this article, standard-error bars are included. Here, however, they are approximately the size of the curve symbols. p = proportion correct; RMSE = root-mean-square error. Number in parentheses indicates the number of free parameters.

Experiment 2: Testing the Independent-Sampling Model

Consider a condition in which a stimulus is displayed, followed by a blank interstimulus interval (ISI) for 250 ms, followed by the *exact same* stimulus again. Compare this condition to two control conditions, one of which consists of just the first presentation, and the other which consists of just the second. Independent sampling implies that performance for two exposures of a stimulus, presented within a single trial and separated by some ISI, should be equal to the probabilistic sum of performance for each separate exposure presented alone.

We use probability summation to quantify the notion of independent sampling. Probability summation is incorporated by our theory quite gracefully. Consider any two information sources (e.g., first and second presentations; left eye and right eye; etc.), and suppose that proportions correct based on these two sources are p_1 and p_2, respectively. If p represents proportion correct given both information sources, then probability summation implies the following:

$$p = p_1 + (1.0 - p_1)p_2, \tag{10}$$

or

$$(1.0 - p) = (1.0 - p_1)(1.0 - p_2).$$

Given performance in terms of $P_i = -\ln(1.0 - p_i)$, probability summation can be expressed as

$$P = P_1 + P_2. \tag{11}$$

Thus, probability summation using our transformed variable P is accomplished by simply adding individual performance values to obtain the probabilistically summed performance value.

In Experiment 2, we explicitly tested the independent-sampling assumption, using the methodology just described. In the *critical condition,* stimulus presentation was split into two parts: A 50-ms initial stimulus presentation was followed by a 250-ms ISI and then by a variable-duration second stimulus presentation. Performance in the critical condition was compared with two control conditions: The first consisted of just the 50-ms initial presentation, whereas the second consisted of just the variable-duration second presentation. The independent-sampling assumption makes a simple prediction: For each second-duration value, performance in the two control conditions sum to performance in the critical condition.

Table 3
Linear Regression Parameters for Experiment 1

Observer	L	c_r	r^2	Asymptote (Y)
AB	12	34.6	0.970	0.990
DA	24	63.8	0.989	0.859
KC	24	50.0	0.958	0.925
TB	17	85.5	0.979	0.931
Average	16.6	53.4	0.988	0.892

Note. L = liftoff. Average fits represent regression parameters derived from fits on averaged data.

Table 4
Linear Regression Parameters for Experiment 2

	Dual-presentation condition			Second presentation condition			Asymptote (Y)	First only (P)
Observer	ReL	c_r	r^2	L	c_r	r^2		
AB	22	53.2	0.999	15	34.9	0.991	0.991	0.911
DA	28	33.2	0.959	13	49.8	0.792	0.804	0.600
KC	28	55.9	0.976	32	41.9	0.973	0.848	0.498
TB	28	70.2	0.977	27	61.2	0.991	0.896	0.339
Average (DA, KC, TB)	28	20.7	0.995	24	20.1	0.951	0.848	0.481

Note. ReL = re-liftoff; L = liftoff. Average fits represent regression parameters derived from fits on averaged data.

Sensory Independence

To use the two control conditions in this probability summation comparison, the critical condition must meet one additional criterion. The two presentations in the critical condition must remain independent at the sensory level; that is, the $a(t)$ curve from the first presentation must not overlap the $a(t)$ curve from the second presentation. To see why this is so, recall first that performance, P, is proportional to $A_\theta(\infty)$, the total above-threshold area under the sensory response function. Thus, the additivity in P (discussed earlier) can be viewed equivalently in terms of additivity of $A_\theta(\infty)$.

Suppose that the presumed threshold were zero. Then by any linear model, $A_\theta(\infty)$, which would equal $A(\infty)$, would be proportional to $F(\infty)$, the total area under the stimulus contrast function. Because the summed $F(\infty)$s for two control conditions equals $F(\infty)$ for the critical condition, additivity would hold in terms of both $A_\theta(\infty)$ and P.

However, this chain of logic breaks down when threshold is greater than zero. Now additivity between the two control conditions requires a long enough ISI separating the first and second critical-condition presentations that the sensory response engendered by the first presentation has essentially decayed to zero before the second presentation's onset. Figure 6 demonstrations this assertion by using one real and one hypothetical condition. Both panels depict a critical condition in which a 50-ms first presentation is followed by a 100-ms second presentation. In the top panel, which represents an actual condition used in Experiment 2, the ISI is long (250 ms) and the two $a(t)$ curves do not overlap, thereby satisfying sensory above-threshold area (0.39) and the second-presentation above-threshold area (1.24).

The hypothetical condition represented in the bottom panel contains a shorter ISI (15 ms), causing the first- and second-presentation $a(t)$ curves to overlap, thereby violating sensory independence (note that the dashed lines show the individual $a(t)$ functions engendered by the first and second presentations separately). The summation process leaves more area above the sensory threshold (1.81) than do the summed control conditions (1.63).

Because it is critical to much of the logic we present, we risk redundancy here to reiterate the effect of a nonzero threshold on additivity. The top and bottom panels of Figure 6 both show $a(t)$ functions engendered by a 50-ms presentation followed by a 100-ms presentation; the only difference between the panels is the ISI that separates the presentations. The theory's linear nature dictates that the *total areas* (i.e., the above-zero areas) are the same (4.95, as it happens) for both panels, even though the *shapes* are obviously quite different. However, when an above-zero threshold is introduced, the rules change. Because higher parts of the function are more likely to exceed threshold, the height of the summed function—and, ipso facto, the shape of the function—influences above-threshold area. When the first- and second-presentation $a(t)$ functions overlap as they do with a short ISI (bottom panel), the summed-function height increases, and more of the function falls above threshold.

Accordingly, to maintain sensory independence in Experiment 2, the gap between the first and second presentations of the control condition was always 250 ms.

Method

The methodology, stimuli, and equipment were very similar to that of Experiment 1. The main change was that the single-presentation con-

Table 5
Linear Regression Parameters for Experiment 4

	Monoptic			Gap monoptic			Asymptote (Y)
Observer	L	c_r	r^2	L	c_r	r^2	
GW	60	180.4	0.992	93	252.8	0.990	0.762
TB	66	112.7	0.996	99	170.8	0.958	0.860
TK	61	128.0	0.993	95	152.9	0.959	0.993
Average	65	117.0	0.999	108	146.0	0.991	0.827

Note. L = liftoff. Average fits represent regression parameters derived from fits on averaged data.

Table 6

Linear Regression Parameters for Experiment 5

Observer	Rectangular function			Ramp on			Ramp off			Asymptote (Y)
	L	c_r	r^2	L	c_r	r^2	L	c_r	r^2	
CW	34	123.6	0.989	42	60.0	0.971	35	75.2	0.962	0.995
LM	44	111.8	0.973	44	56.7	0.970	44	56.6	0.993	0.993
TB	30	160.7	0.940	42	59.3	0.989	37	87.7	0.990	0.998
SB	42	192.0	0.851	37	99.4	0.992	43	85.4	0.993	0.997
Average (CW, LM, TB)	37	126.3	0.997	43	60.1	0.990	39	70.2	0.995	0.981

Note. L = liftoff. Average fits represent regression parameters derived from fits on averaged data.

dition of Experiment 1 was compared with a new, double-presentation condition.

Observers. Four observers participated in the experiment: the first author (TB), one undergraduate student (KC), and two graduate students (AB and DA). All observers were highly practiced, having participated in a minimum of 1,700 trials before beginning this experiment.

Stimuli and apparatus. Stimulus presentation was the same as in Experiment 1, save for one difference. First, during the initial warning tone for each trial, a fixation point appeared briefly over the permanent fixation point mounted to the projector screen. The fixation point consisted of a white square superimposed over the adapting field. Fixation-point luminance and contrast are provided as part of Table 2.

Design. Three different conditions were used, each involving six durations. These conditions are schematized in Figure 7.

1. In the *dual-presentation critical condition*, the digit array was presented for 50 ms, followed by a 250-ms adaptation-field exposure, followed by a second presentation of the exact same digits. This second presentation ranged in duration from 17 ms to 100 ms and was followed immediately by an adapting field for 300 ms.

2. The *first-only control condition* consisted of the 50-ms first presentation, presented alone, followed by a variable-duration adaptation field before the onset of the warning tone. The adaptation-field duration ranged from 250 ms to 333 ms and varied to remain as close as possible to the conditions of the dual-presentation condition. This first-only condition provides a baseline performance level for the 50 ms first presentation in the dual-presentation condition.

3. The *second-only control condition* consisted of just the variable-duration second stimulus, presented alone for a duration that ranged from 17 ms to 100 ms. This presentation was preceded by 300 ms of the adapting field. We call this condition the second-stimulus condition, and it provides a performance curve similar to that generated in Experiment 1.

As is graphically evident in Figure 7, combining the two control conditions yields the dual-presentation stimulus condition.

Procedure. Each observer participated in twenty-four 72-trial blocks, which provided 96 observations per condition per observer. Within each block, the 18 stimulus conditions were randomly intermingled over the 72 trials. The sequence of events both within a block and within a trial were the same as in Experiment 1 except for the minor changes described earlier. As in Experiment 1, the adaptation field was on continuously.

Results

The results of Experiment 2 are shown in Figure 8; recall also that regression parameters are in Tables 3–6 and theory parameters are in Table 1. Three observers (DA, KC, and TB) showed similar data patterns, and accordingly we present their

results averaged together (top panel). A fourth observer (AB) had a somewhat different data pattern; her data are shown in the bottom panel and are discussed below. In all cases, the solid lines represent the best fit of the theory.
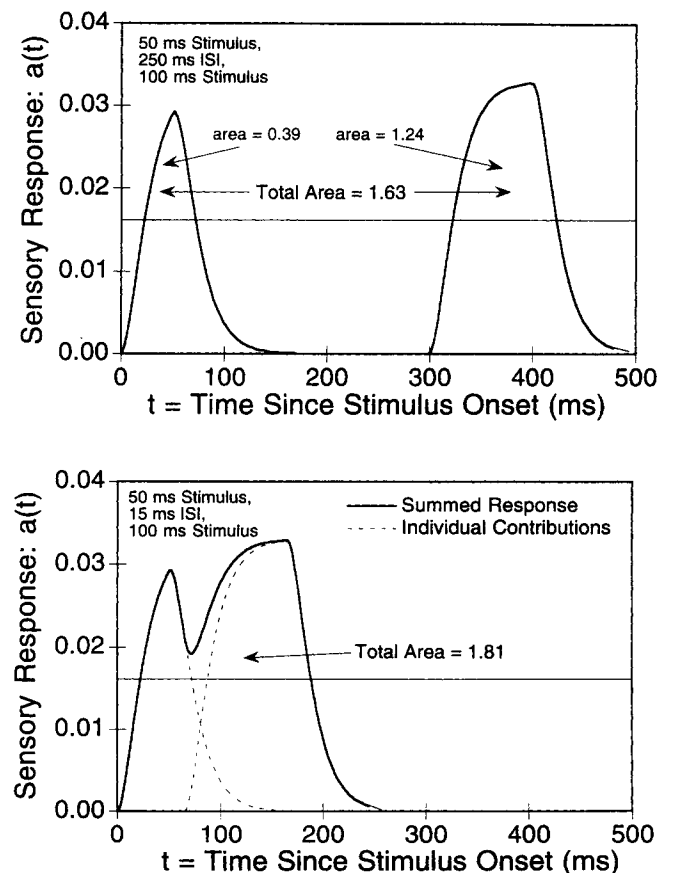

Figure 6. Top panel: $a(t)$ curves for the same two presentations separated by a 250-ms interstimulus interval (ISI). Although total area is the same for both curves, above-threshold area is greater for the top-panel curve. Bottom panel: $a(t)$ curve produced by a 50-ms and 100-ms presentation separated by a 15-ms ISI. Dashed lines show the individual contributions of the two separate stimuli, whereas solid lines show the summed response.
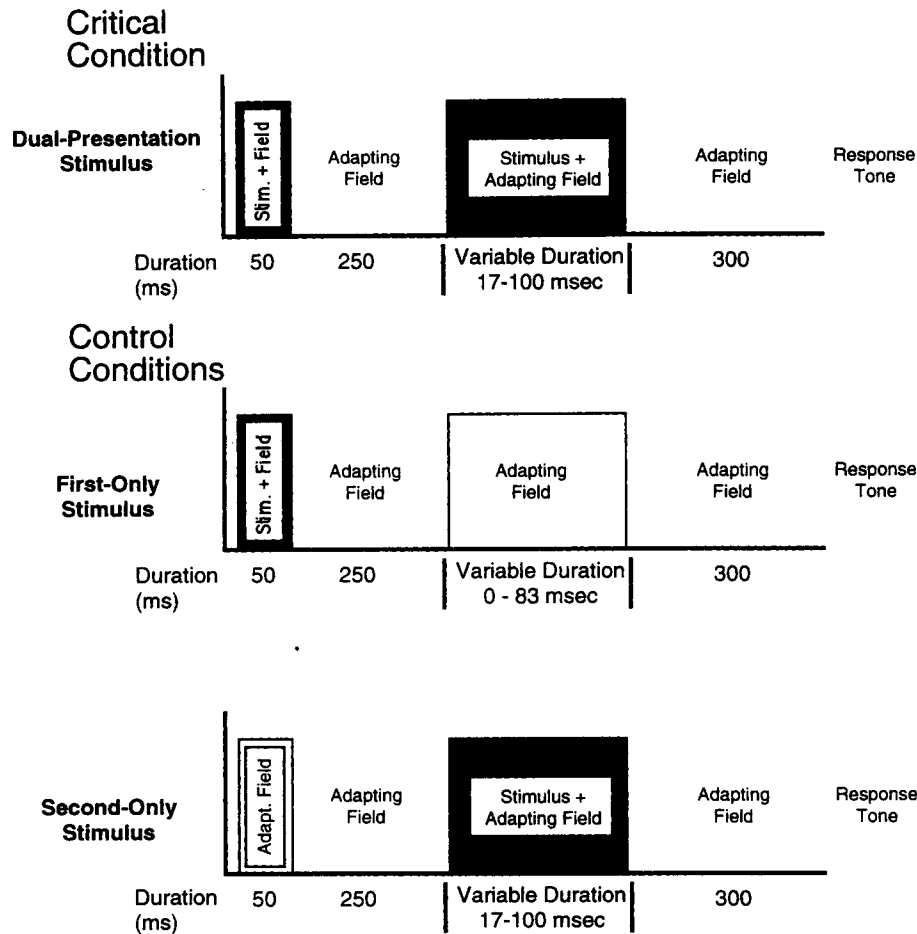
Critical Condition

Dual-Presentation Stimulus

| Stim. + Field | Adapting Field | Stimulus + Adapting Field | Adapting Field | Response Tone |

Duration (ms) 50 250 | Variable Duration 17-100 msec | 300

Control Conditions

First-Only Stimulus

| Stim. + Field | Adapting Field | Adapting Field | Adapting Field | Response Tone |

Duration (ms) 50 250 | Variable Duration 0 - 83 msec | 300

Second-Only Stimulus

| Adapt. Field | Adapting Field | Stimulus + Adapting Field | Adapting Field | Response Tone |

Duration (ms) 50 250 | Variable Duration 17-100 msec | 300

Figure 7. Display conditions for Experiment 2. Combining the two control conditions (bottom two panels) equals the critical, dual-presentation condition (top panel). Stim. = stimulus; Adapt. = adapting.

Observers DA, KC, and TB. As anticipated, the first-only condition data (solid squares) were essentially flat: Adding additional time before the response cue did not affect performance.

The dual-presentation condition (open circles) is conceptually the sum of the two control conditions: both observed and predicted performance reflect this in the sense that the dual-presentation condition curve is shifted upward relative to the control curve by an amount approximately equal to the 50-ms, first-only performance. Correspondingly, the theoretical prediction for the dual-presentation condition (upper solid line) is the sum of the two predicted control-condition curves (bottom two solid lines).

In short, for these three observers, the independent-sampling model, along with the overall theory, is strongly confirmed.

Observer AB. Observer AB differs from the other three observers in an important respect: The slope of the dual-presentation curve is shallower than that of the second-only control curve. This finding is inconsistent with the simple version of our theory. The theory could be modified in one of two ways to account for AB's data pattern. First, it might be assumed that

the first presentation caused a slower feature-sampling rate during the second presentation. Second, it might be assumed that the first presentation caused a higher sensory threshold on the second presentation. Either of these modifications allows the theory to account reasonably well for AB's data, although only the higher second-threshold accounts for certain, somewhat minute aspects of the data. As an illustration of how the increased-threshold assumption would account for the data, we estimated two thresholds for AB's dual-presentation condition: one for the first presentation and one for the second presentation. The two control conditions both used the first threshold. These two threshold values are included in Table 1, and the data fits shown at the bottom of Figure 8 result.

Using an additional piece of notation, we can directly test the proposition that AB's differing slopes result from a threshold change from the first to the second presentation. Earlier we described the liftoff parameter as the maximum stimulus duration that gives chance performance. For the second-only presentation, this stimulus duration represents the minimum duration at which an observer can begin acquiring information from the second-only stimulus. An analogous parameter exists for the
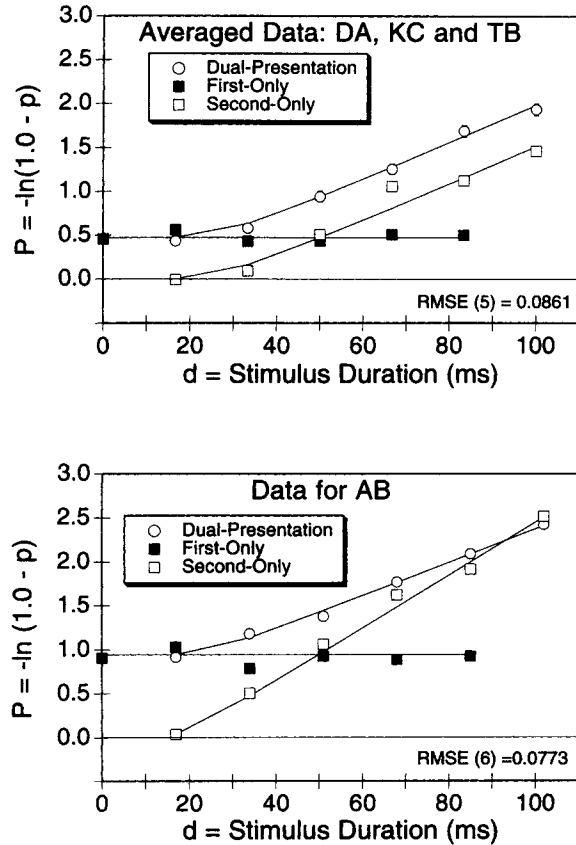
Figure 8. Experiment 2 results. Top panel: averaged data and model fit for three observers (DA, KC, and TB). Bottom panel: results and model fit from Observer AB, assuming a higher second-presentation Θ. p = proportion correct; $P = -\ln(1.0 - p)$; RMSE = root-mean-square error. Number in parentheses indicates the number of free parameters.

dual-presentation condition, which is the point at which the dual-presentation curve rises above the first-only baseline. This point, which we call *re-liftoff*, or *ReL*, represents the minimum second-only duration at which an observer could acquire information from just the second presentation in the dual-presentation condition. If independent sampling holds, then liftoff and re-liftoff should be equal. As Table 4 shows, re-liftoff is longer than liftoff for two observers, one of whom is AB.[7]

This slightly longer re-liftoff is a direct prediction of the slightly higher second threshold. As threshold increases, the shortest duration that elicits an above-chance response increases. Both AB's data and the theoretical predictions generated from a dual-threshold model to her data have a difference between liftoff and re-liftoff of 6–7 ms. Such a difference would not be predicted by a model that assumed a slower feature-sampling rate during the second presentation.

Discussion

For three of the four observers, we could predict dual-presentation performance based on the sum of the first and second presentation performances. The 250-ms ISI is long enough to

allow the first $a(t)$ function to decay to zero and avoid sensory interactions, and the first presentation does not markedly influence the processing of the second presentation.

The general theory accounts well for these data. Two implications emerge from the theoretical fit. First, the two events of the dual-presentation stimulus condition appear to be processed independently. Second, our current conceptualization of the sensory threshold is tenable. This conceptualization explains not only the above-zero d-intercept (liftoff) for the second-only stimulus but also the above-zero first-only intercept (re-liftoff) for the dual-presentation stimulus.

With a small modification, the theory also accounts for differences between observers. For three observers, independent sampling was supported, and the theory fit the observed data patterns. Independent sampling failed for Observer AB, however. The theory can account for this failure with the addition of a second, higher sensory threshold for the processing of a second presentation in a dual-presentation condition. From this we conclude that although the general theory accounts for the majority of observers, the theory must remain flexible enough to account for individual differences.

Experiment 3: ISI Manipulation and Sensory Independence Failure

In Experiment 2, we focused on the independent-sampling assumption. To test this assumption as directly as possible, we used a 250-ms ISI to avoid sensory interactions between the first and second presentations in the critical condition. In discussing the logic underlying this procedure (see Figure 6) we pointed out that, given a nonzero threshold, the shorter the ISI between the first and second presentations in the critical condition, the greater the value of $A_\Theta(\infty)$; thus, the higher performance should be.

In Experiment 3, we test this assertion directly, thereby focusing on the critical role of the sensory threshold. The design of Experiment 3 was simple: On all trials, there was a 30-ms presentation followed by a 45-ms presentation, separated by a variable-duration ISI that ranged from 0 ms to 105 ms. This design is similar to a detection experiment by Roufs and Blommaert (1981) that was designed to describe the shape of the impulse response function.

Method

Stimulus presentation (and response collection) in Experiment 3 were carried out on the Macintosh II computer. Care was taken to match contrast, luminance, and stimulus conditions as much as possible to those of Experiments 1 and 2. One major difference between the two displays was the display medium. Whereas the tachistoscopic shutter afforded an almost perfect rectangular wave, the computer monitor displays an image by "refreshing" the phosphor screen 67 times a second (see Busey, 1992). Appendix D details the implications for our theory of using a computer monitor as a display device. To summarize, the linear-filter model virtually equates steady displays with flickering displays, provided the flicker occurs faster than some threshold rate. The flicker

---

[7] Interpreting the liftoff and re-liftoff values obtained from the regression analyses is somewhat problematical, however, as, by the theory, the performance curves are slightly nonlinear at short durations.
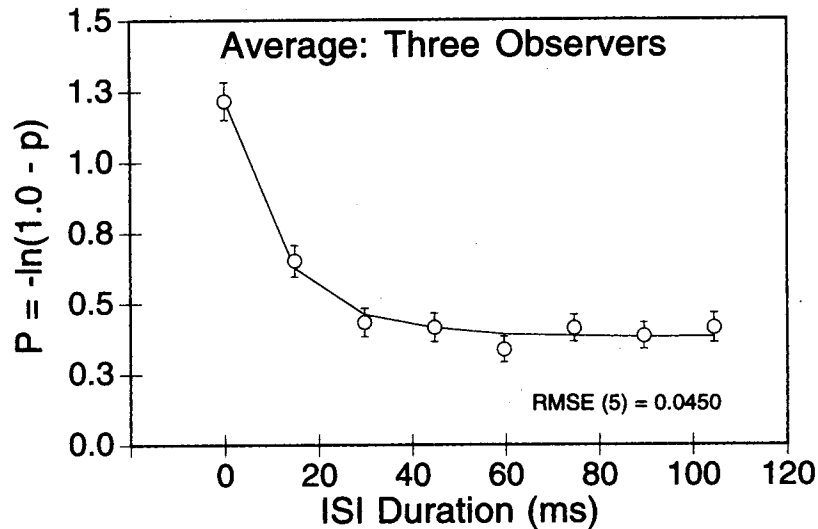
Figure 9. Experiment 3 results. RMSE = root-mean-square error; ISI = interstimulus interval; p = proportion correct; $P = -\ln(1.0 - p)$. Number in parentheses indicates the number of free parameters.

rate (67 Hz) of the display device exceeds the threshold rate necessary for flicker-free displays.

Observers. Three observers, the first author (TB), Observer DA from Experiment 2, and an undergraduate student (TK) participated in Experiment 3. All observers had participated in a minimum of 1,000 trials before participating in Experiment 3.

Stimuli and apparatus. Experiment 3 was controlled by a Macintosh II computer, and stimuli were displayed on an Apple Monochrome monitor. Observers sat approximately 57 cm away from the screen in a dimly lit room and used the computer keypad to respond.

The same font (Times-Roman) used to create the stimulus slides for previous experiments was scaled to present the same image size in Experiment 4. Each letter subtended a visual angle of 0.50° vertically and 0.40° horizontally, separated by 0.75° vertically and 0.40° horizontally. Table 2 lists the luminances and contrasts. Darker letters on a lighter background were used to avoid phosphor-decay problems. The experiment was programmed by using a timing and display package described by Ames and Palmer (1992).

Design. In all trials, a 30-ms stimulus presentation was followed by a variable-duration, blank ISI, which was followed by a 45-ms stimulus presentation. ISI ranged from zero ms[8] to 105 ms in eight 15-ms increments. Observers completed eight 72-trial blocks, which provided 72 observations per condition per observer.

Procedure. The procedure was similar to that of Experiments 1 and 2. A trial began with a 250-ms warning tone that reminded observers of the current to-be-reported row for the current block. Following the tone, the digit array was displayed in its appropriate temporal configuration, followed 375 ms later by a response tone. The observer then tried to type in the digits in their proper order, guessing when necessary.

Results and Discussion

Figure 9 shows the mean results for three observers, along with the corresponding model fits. Longer ISIs produced a performance decrement. The decrement is entirely consistent with the proposition, central to the theory, that there is a sensory threshold, and that performance is determined by $A_\theta(\infty)$, the above-threshold area.

As usual, the theoretical parameters are provided in Table 1, and the solid line in Figure 9 shows the best theoretical fit. The fit is quite good, with an RMSE of 0.045.

Experiment 4: One Versus Two Onsets

In Experiment 4, we systematically investigated a question that was raised in both Experiments 2 and 3: What is the effect of displaying a stimulus that has a single onset versus two distinct onsets? In particular, in Experiment 4, we examined stimuli presented for *d* ms that were either shown all at once (which we term a *single-onset condition*) or that were shown in the form of 2 (*d*/2)-ms halves separated by a 250-ms ISI (which we term a *dual-onset condition*).

Kowler and Sperling (1980) reported a similarly motivated experiment that was carried out to test the possibility that stimulus onset may play an important role in visual information processing. On the basis of data showing strong positive effects of number of eye fixations on visual processing (Loftus, 1972) as well as data showing stimulus onset to be more important than stimulus duration (Sperling, Budiansky, Spivak, & Johnson, 1971; Sperling & Melchner, 1978), Kowler and Sperling reasoned that a dual-onset condition would produce better memory performance relative to a single-onset condition. However, Kowler and Sperling found no effect of stimulus onset at all, positive or negative.

As argued earlier (see Figure 6), our theory predicts that, as long as there is a greater-than-zero sensory threshold, the single-onset condition will produce superior performance to the dual-onset condition. Although when taken at face value, the Kowler

---

[8] Because we are using a computer monitor, the smallest ISI was equal to the refresh rate, or 15 ms. However, given Appendix C and the inability of the visual system to temporally resolve the flicker rate of the monitor, we assert that the effective ISI was zero.

and Sperling (1980) data appear at odds with this prediction, their design carries some interpretational problems. First, their stimuli were of relatively high contrast, which could render the hypothesized threshold small and irrelevant compared with the total magnitude of the $a(t)$ function. Second, and more important, Kowler and Sperling's single- and dual-onset conditions involved presentations that entailed different amounts of total presentation time. In their single-onset condition, the stimulus was on continuously for d ms, whereas in the dual-onset condition, the stimulus was turned on and off quickly, then, following a $d/2$ ms stimulus onset asynchrony, was turned on and off quickly a second time. Accordingly, the physical stimulus display time was less in the dual-presentation than in the single-presentation condition. (We do note that although this confound poses interpretational problems in this treatment, the original Kowler and Sperling design is appropriate given their empirical question.)

In Experiment 4, we refer to the single-onset condition as the *monoptic* condition and the dual-onset condition as the *gap-monoptic* condition, as each of the conditions involves stimulus presentation to only a single eye. The reasons for presenting stimuli monoptically will unfold as we discuss extensions of the linear-filter model to binocular combination tasks.

Method

Stimulus presentation and response collection in Experiment 4 were carried out on the Macintosh II computer in conjunction with a mirror stereoscope, which is depicted in Figure 10.

There were two basic conditions in Experiment 4. In the monoptic condition, a single d-ms presentation was displayed to one eye. In the gap-monoptic condition, two $(d/2)$-ms presentations were separated by a 250-ms ISI, shown, again, to only one eye. There were eight exposure durations within each of these two monoptic/gap-monoptic conditions, for a total of 16 conditions.

Observers. Three observers, the first author (TB), one graduate student (TK), and one Dartmouth College faculty member (GW), partici-

pated in Experiment 4. All observers had participated in a minimum of 1,000 trials before the experiment.

Stimuli and apparatus. Experiment 4 was controlled by the Macintosh II computer used in Experiment 3. Stimuli were displayed on an Apple Monochrome monitor in conjunction with a Modified Wheatstone mirror stereoscope (see Blake & Fox, 1973). Observers sat approximately 57 cm away from the screen in a dimly lit room and used the computer keypad to respond. A chin rest positioned the observer's head in front of two mirror arrangements that projected each half of the screen to one of the observer's eyes. An enclosing box eliminated reflections from the screen and assisted with eye fusion.

Considerable effort was taken at the start of each block of trials to fuse the two halves of the monitor into one image without eye strain. Observers adjusted each of the four mirrors until the fixation point fused into a single image. They then adjusted the mirrors to create a cross out of two right angles, each of which was shown only to one eye. Figure 10 shows the experimental apparatus and the image used to adjust the mirrors. To further enhance eye fusion within a session, the background was a uniform gray field, and the fixation point never disappeared.

Design. Stimulus duration ranged from 30 ms to 240 ms in eight 30-ms increments. In the monoptic condition a stimulus was shown once for d ms to either the right or left eye. Eye position was randomized across all conditions. In a corresponding gap-monoptic condition, a stimulus was presented for $d/2$ ms to either the right or left eye, followed by a 250-ms ISI during which the digits were absent, followed by another $d/2$ ms presentation to the same eye. During the presentation to one eye, the other eye saw the blank background field.

Each observer participated in eighteen 72-trial blocks, which provided 81 observations per condition per observer.

Procedure. The procedure was similar to that of Experiments 1 through 3. A trial began with a 250-ms warning tone that reminded observers of the to-be-reported row for the current block. Following the tone, the digit array was displayed in its appropriate temporal configuration, followed 540 ms after stimulus onset by a response tone. The observer then tried to type in the digits in their proper order, guessing when necessary.

Results and Discussion

Figure 11 shows the results averaged over the three observers, with the corresponding model fit. Several aspects of the data
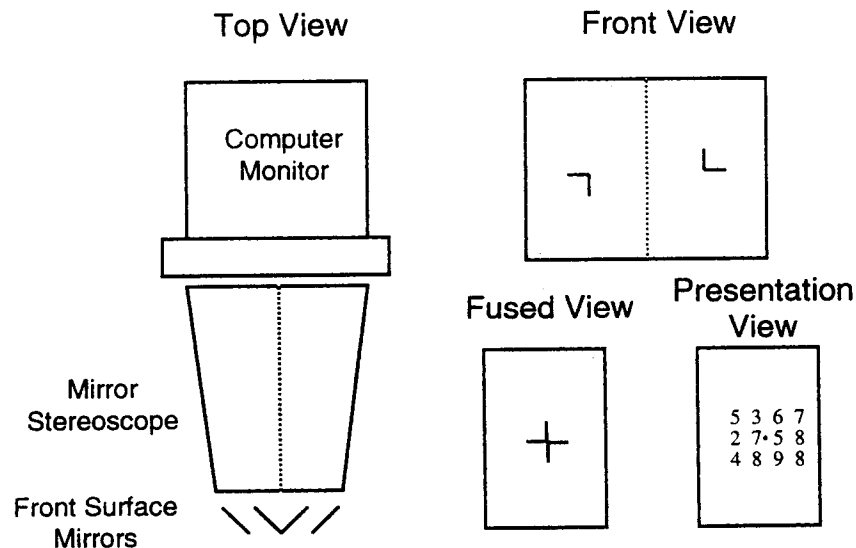

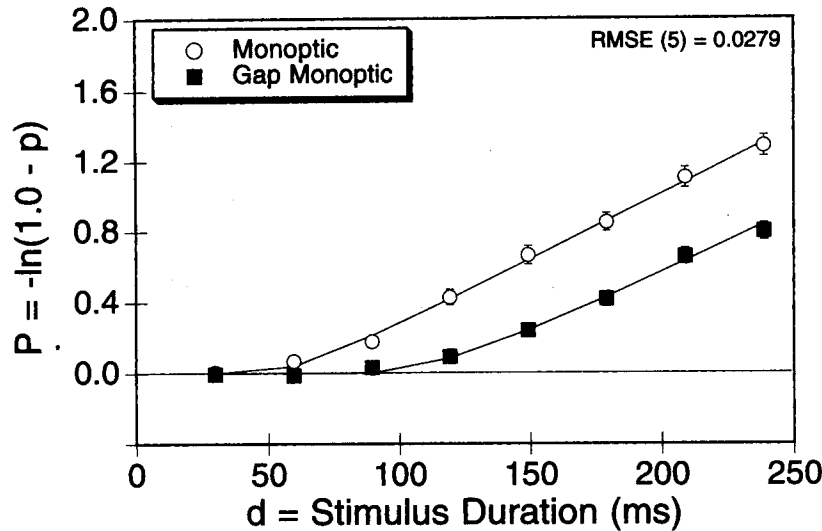*Figure 10.* Display apparatus for Experiment 4.

Figure 11. Experiment 4 results. RMSE = root-mean-square error; p = proportion correct; $P = -\ln(1.0 - p)$. Number in parentheses indicates the number of free parameters.

are notable. First, it is clear that, as in past experiments, both performance curves are highly linear. Second, as the theory predicts and Experiment 3 demonstrated, inserting a gap in the middle of the stimulus presentation substantially reduces performance. This latter finding is superficially counter to that of Kowler and Sperling (1980); however, as we noted earlier, their null effect could be a result of two factors. First, the higher stimulus contrast that they used could have placed their task in a *resource-limited* domain, whereas this experiment, with its much lower contrast, is in a *data-limited* domain (Sperling, 1979). Performance in resource-limited tasks depends on the total stimulus exposure time, with longer stimulus presentations providing more time for stimulus processing such as transfer to more permanent memory. Performance in data-limited tasks depends on the total energy in the stimulus, with longer stimulus presentations providing more above-threshold energy.

The second factor could have involved Kowler and Sperling's (1980) procedure of confounding two onsets with less physical exposure duration: The total contrast energy was different in the two conditions, which makes comparisons difficult within the context of the current data-limited theory.

The theory fits the data well: As indicated in Figure 11, the RMSE is 0.0279. Table 1 summarizes the best fitting parameters for individual observers and for the combined data.

Experiment 5: Ramped Stimuli

In Experiment 5 we dramatically altered the shape of the $f(t)$ function. Experiments 1 through 4 all used rectangular-function presentations in which a stimulus abruptly appeared, remained on for the requisite duration, and then abruptly disappeared. Experiment 5 introduced two new $f(t)$ functions: a ramp-on function in which the stimulus appeared gradually before disappearing abruptly and a ramp-off function in which the stimulus appeared abruptly before disappearing gradually.

Method

The methodology, stimuli, and equipment of Experiment 5 were very similar to that of Experiment 4. The main change was that the rectangular-function presentation condition of Experiment 4 was compared with two new, ramped-presentation conditions. Stimuli were viewed binocularly, without the use of a mirror stereoscope.

Observers. Four observers, the first author (TB), one male graduate student (MB), one female graduate student (LM), and one male undergraduate student (SB), participated in Experiment 5. All observers had participated in a minimum of 1,000 similar trials before the experiment.

Stimuli and apparatus. Experiment 5 was controlled by the Macintosh II computer used in Experiment 4. Stimuli were displayed on an Apple Monochrome monitor. The ramping-on stimuli followed a linear contrast gradient from zero contrast up to a contrast of twice the rectangular-function's contrast. The ramping-off stimuli followed a linear contrast gradient from a contrast of twice the rectangular function's contrast down to zero contrast. This equalized the amount of contrast energy between the three types of stimuli for a given duration. Because the display device was a cathode ray tube (CRT) the ramping functions were sampled at 15-ms intervals to approximate a true ramping function.[9] Subsequent modeling procedures took this approximation into account. Figure 12 provides a graphical summary of the three types of stimuli.

Design. Stimulus duration ranged from 30 ms to 145 ms in eight 15-ms increments. Each observer participated in twenty-four 72-trial blocks, which provided 72 observations per condition per observer.

[9] Simulating ramping functions with a CRT has the unfortunate result of leaving the first (for ramp-on stimuli) or last (for ramp-off stimuli) refreshes of the stimulus duration at zero contrast, effectively reducing the duration of these stimuli by one screen refresh (15 ms) in relationship to the same-duration step function. To account for this, we added a small correction factor of 0.43% contrast units to the first or last refreshes and adjusted the rest of the contrast steps in each presentation to account for this. Care was taken to equalize the amount of contrast energy between the three types of stimuli for a given stimulus duration.
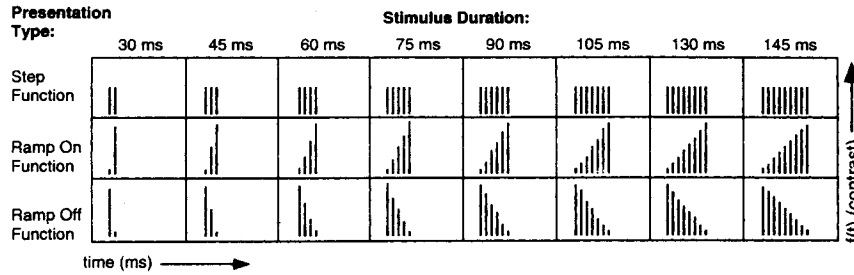
Figure 12. Experiment 5 design. Each bar represents the contrast level at each screen refresh. $f(t)$ = stimulus contrast function.

Procedure. The procedure was similar to that of Experiments 1 through 4. A trial began with a 250-ms warning tone that reminded observers of the to-be-reported row for the current block. Following the tone, the digit array was displayed in its appropriate temporal configuration, followed 540 ms after stimulus onset by a response tone. The observer then tried to type in the digits in their proper order, guessing when necessary.

Results and Discussion

Figure 13 shows the results averaged over the four observers, with the corresponding model fit. Several aspects of the data are notable. First, it is clear that, as in past experiments, both performance curves are highly linear.

The theory provides a good fit to the data: As indicated in Figure 13, the RMSE is 0.0584. The one significant violation occurs at the three longest stimulus durations for the ramp-on and ramp-off conditions. The data show a slight superiority for the ramp-on stimuli. We investigated the possibility that this result could be accounted for with a multiphasic, impulse response function, but this extension to the model was unsuccessful. Table 1 summarizes the best fitting parameters for individual observers and for the combined data.

Experiment 6: Detection and Identification—Two Tasks Sharing a Common Percept?

Our theory successfully accounts for data in character identification tasks. Similar linear systems models have been applied by Watson (1986) and others to the detection of sinusoidal patches in the study of low-contrast detection. Experiment 6 tests whether our theory can simultaneously account for data in both domains.

In Experiment 6 we presented a single digit in a two-alternative, forced-choice (2AFC) paradigm. Two responses were recorded on each trial: one indicating which of two temporal intervals contained the digit, and the other indicating the digit's identity. We then determined the contrast that yielded 81% correct for both the detection and identification tasks. Separate contrast thresholds were estimated for each task. These contrast thresholds were then modeled by our theory to see if the same parameters could account for performance in both tasks.

Method

Stimulus presentation and response collection in Experiment 6 were similar to that of Experiment 5. Contrast thresholds for eight stimulus

durations were measured by using an adaptive threshold determination procedure (Quest) developed by Watson and Pelli (1983) in a two-temporal-alternative, forced-choice procedure. A single digit (either a 2 or a 5) was presented for one of eight stimulus durations: 15, 30, 45, 60, 75, 90, 105, or 145 ms. Stimulus presentation occurred in one of two temporal intervals designated by warning tones. Observers made two responses: the first indicating which of the two temporal intervals contained the digit, and the second indicating the digit's identity.

Observers. Three observers, the first author (TB), one male graduate student (MB), and one female graduate student (LM), participated in Experiment 6. All observers had participated in a minimum of 1,000 similar trials before the experiment.

Stimuli and apparatus. Experiment 6 was controlled by the Macintosh II computer used in Experiment 5. The stimuli were the same size as each digit of the Experiments 1–5 stimuli. The digit always appeared in the same location, centered vertically on the screen and 0.27° below the fixation point. Stimulus contrast was varied on a trial-by-trial basis. The stimulus contrast function was rectangular.

Design. Each observer participated in twenty 64-trial blocks, which provided 160 observations per condition per observer.

Procedure. The procedure on a single trial was as follows. Observers heard a warning tone, signifying the start of the first temporal interval. After a 30-ms delay, the stimulus appeared with a 50% probability. Seven hundred milliseconds after the end of the first warning tone, a second warning tone indicated the end of the first temporal interval and the beginning of the second interval. The second interval was the same duration as the first, again contained the stimulus with a 50% probability, and was terminated by a response tone.
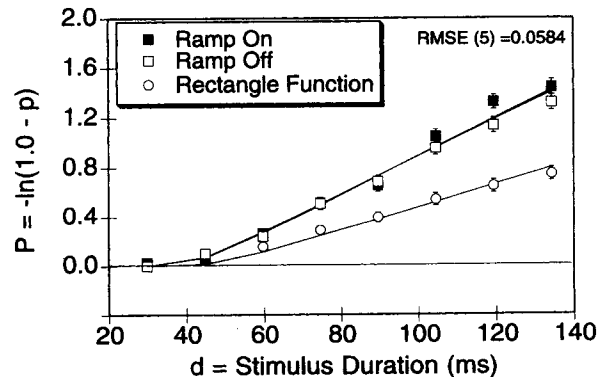

Figure 13. Experiment 5 results. RMSE = root-mean-square error; p = proportion correct; $P = -\ln(1.0 - p)$. Number in parentheses indicates the number of free parameters.

Observers made two responses, one indicating the temporal interval that contained the digit, guessing if necessary, and a second response indicating the digit's identity, again guessing if necessary. Each response was followed by feedback in the form of a high or low tone.

Results and Discussion

Figure 14 shows the individual results for the three observers, with the corresponding theory fits, plotted on log–log axes for reasons described later. Several aspects of the data are notable. First, for two of the three observers, the identification task contrast thresholds were greater than the corresponding detection task thresholds; that is, these observers required more contrast in the identification task than in the detection task to achieve a given performance level. To fit these data requires an additional assumption: Although both tasks have the same amount of information available in the sensory response function, they make use of that information at different rates. Thus we ascribe different sampling rates, c_s, to each task (although we assume them to share common τ and Θ parameters). Given this additional assumption, our theory provides an excellent fit to the Experiment 6 data, as demonstrated by the data and model fits in Figure 14.

Performance in the two tasks was nearly identical for subject LM, as demonstrated by the lower panel of Figure 14. Our original theory with a single c_s provides an adequate description of the data, although adding the additional c_s for the identification task data decreases the RMSE to 0.00202.

These detection data are quite similar to data reported in experiments designed to determine the critical duration, defined as the duration at which the slope first changes on a log–log plot such as those of Figure 14. The estimated critical durations for all three subjects are in the range of 45–60 ms and are similar to those reported by others. The contrast thresholds below the critical duration have a slope of approximately −1.0 when plotted on log–log coordinates, as predicted by Blochs Law. (The Figure 14 plots have lines with −1.0 slopes for comparison.) The contrast thresholds for longer durations have a slope that is greater than −1.0 but less that zero, which corresponds to predictions made by Watson (1986, Figure 6.13) for a probability summation model with a similar monophasic impulse response function.

The linear-filter model, with a single τ and Θ, provides an adequate description of both detection and identification performance. Thus, it appears that the two tasks can be construed as sharing a common sensory input in the form of the sensory response function and differing only in their respective information-sampling rates.

General Discussion

Our major goal in this article was to develop and test a theory of early visual processes and their relation to performance in a simple memory task. We discuss the success with which this goal has been achieved, examine several implications for low-contrast detection research, and preview how the theory may be extended to account for performance in binocular summation tasks.
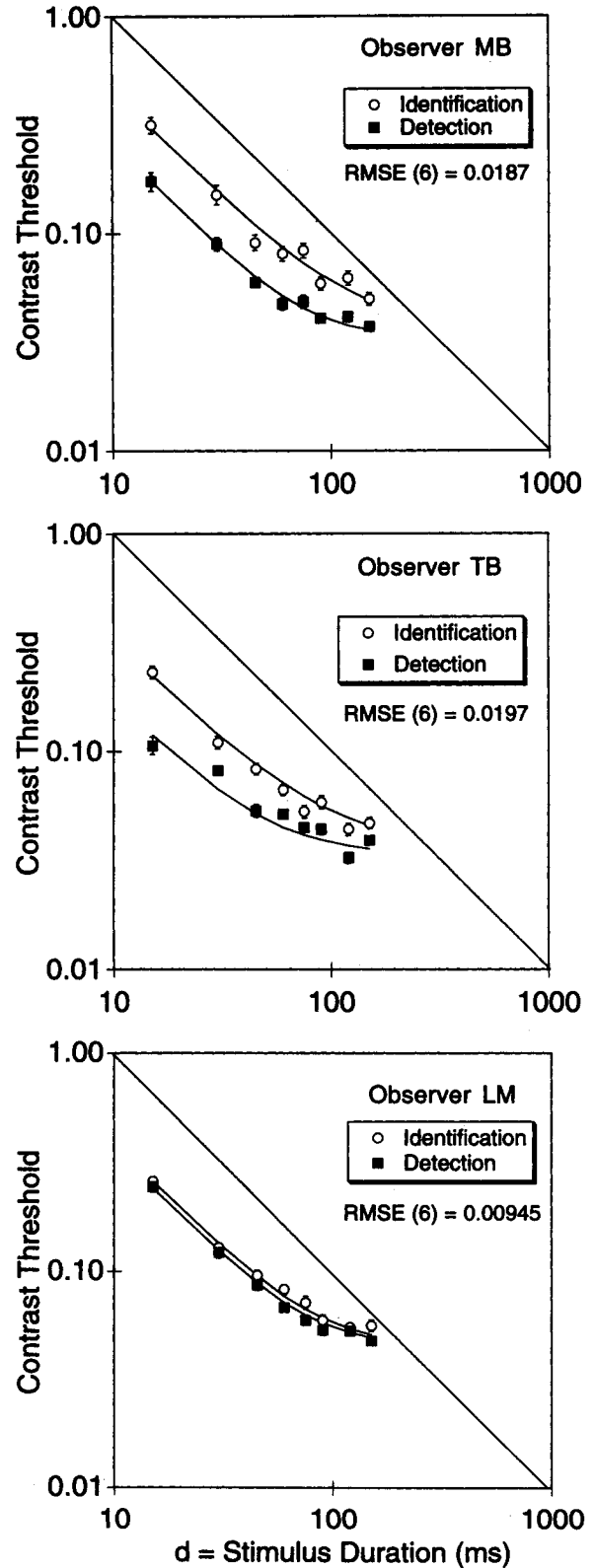

Figure 14. Experiment 6 results. RMSE = root-mean-square error. Number in parentheses indicates the number of free parameters.

Theory

The theory that we have presented has roots in two past traditions. The proposition that the initial visual stages constitute a linear, low-pass, temporal filter has proved fruitful in many domains of vision science (Watson, 1986, provides a comprehensive review of application of such models to temporal processes). The proposition that information acquisition consists of random feature sampling has proved similarly fruitful in various domains of perception, learning, and cognition. The uniting of these two traditions has resulted in a theory that allows precise and simple predictions based primarily on physically measurable stimulus waveform characteristics. The theory has proved capable of handling a reasonably wide range of data within the limited domain of the digit-recall paradigm that we have used here. Other work (Loftus et al., 1993; Loftus & Ruthruff, 1994) has been shown to be capable of handling other variants of the digit-recall paradigm (involving masked stimuli and stimulus intensity manipulation) and of serving as a basis for long-term recognition of complex naturalistic pictures.

The one theoretical component that prevents us from relating memory performance directly to the stimulus waveform is the sensory threshold, which is an unobservable construct that must be inferred from the data. However, the existence of such thresholds has long been recognized and has been central in the application of other linear-filter theories to low-level sensory data. For instance, detection of near-threshold stimuli is often explained by assuming that detection occurs when some threshold is exceeded by a stochastically varying response signal (Watson, 1979, 1986). In our Experiments 1–3, the findings of (a) an above-zero liftoff, (b) the dependence of performance on first- or second-presentation ISI, and (c) the superiority of single-onset relative to dual-onset stimuli are all readily explained by the assumption of a sensory threshold.

One current deficiency in the conceptualization of the sensory threshold Θ is its deterministic nature: If a stimulus's $a(t)$ curve fails to rise above Θ, performance will always be at chance for that stimulus. However, there is clear empirical evidence that performance rises with a sigmoidal function rather than a step function as contrast is increased. Fortunately for our current conceptualization, this sigmoidal function is quite steep, rising to asymptote over only a one-octave contrast range. Thus, our conceptualization serves as a reasonable approximation (Denis Pelli, personal communication, September 20, 1993).

Implications for Low-Contrast Detection

Models of low-contrast detection typically characterize the energy of a stimulus as the squared signal function, integrated over space, spatial frequencies, time, or a combination of all three (e.g., Carlson & Klopfenstein, 1985; Legge, Kersten, & Burgess, 1987). In general, such contrast discrimination experiments empirically derive an estimate of threshold contrast energy following the seminal pedestal experiment of Nachmias and Sansbury (1974): Two gratings are presented in a temporal 2AFC procedure, one of contrast C and the other of contrast $C + \Delta C$. The increment contrast ΔC is then varied to find a value that produces a criterion (e.g., 75%) level of performance. This procedure is then repeated for gratings of different contrasts to

provide a contrast discrimination function relating contrast C to change in contrast ΔC.

These models typically have not considered the temporal characteristics of the stimulus signal, instead focusing on the level of internal noise (Legge et al., 1987; Pelli, 1981) or the effects of different spatial frequencies (Graham, 1977). We have found it useful to include a definition of the temporal response of the visual system to a stimulus within our theory, and the theory presented in this article provides a framework for such a conceptualization. Any low-contrast detection model that assumes that performance is related to the extent to which a response exceeds some threshold (e.g., the threshold transducer model of Foley & Legge, 1981) requires computation of some threshold-dependent response. However, this threshold-dependent response depends on the temporal configuration of the original stimulus, which our theory can provide.

The current data do not imply clear conclusions about the nature of the *transducer function* that relates the physical stimulus energy to the magnitude of the visual system's response. Our theory proposes a linear transducer function, which is contradicted by some (e.g., Legge, 1980; Gottesman et al., 1981) but not all empirical findings (e.g., Graham, 1977; Sachs et al., 1971). A square-law transducer function was applied to some of our data but produced distinctly poorer fits. Loftus and Ruthruff (1994) varied contrast as well as stimulus duration and concluded that a linear transducer function could account for performance for contrast up to 21%. We conclude that although our theory provides a good description of the temporal interactions of the stimulus, it does not discriminate very well between various alternative transducer functions. Recent detection task work with externally added noise has successfully distinguished between the level of an observer's internal noise and the observer's sampling efficiency in the presence of that noise. These models assume that the stimulus is sampled in the presence of intrinsic internal noise but at a level that is less than optimal. Either a higher internal noise level or poorer sampling efficiency will result in higher observed thresholds. However, the addition of external noise to the stimulus distinguishes between these two factors, leading to the conclusion that the level of internal noise, rather than changes in sampling efficiency, is responsible for contrast discrimination performance (Legge et al., 1987). The sensory threshold, Θ, will be proportional to the square root of the sum of the unknown internal noise and the added external noise (Denis Pelli, personal communication, October 13, 1993). Added-noise experiments may help refine the current conceptualization of the sensory threshold and the sensory response function $a(t)$. We are currently undertaking such work.

Extending the Theory to Binocular Combination

The theory presented here can be readily generalized to the domain of binocular combination. We normally view the world binocularly, and somehow the information from each eye combines to give us a cyclopean view of the world. The theory stresses the temporal aspects of visual perception and can be quite useful in examining the temporal interactions between the sensory information coming from each eye. A forthcoming article will describe research carried out on this topic; we briefly preview it here.

Additional Assumptions

To account for the added complexity of the binocular combination domain, three additional assumptions are required. For the binocularly and monocularly viewed stimuli of Experiments 1 through 6, the current theory is a special case of these assumptions.

We assume that each eye acts as a linear temporal filter on the stimulus to produce individual $a(t)$ functions. These functions are termed the *peripheral* $a(t)$ *functions* and are designated by the notation $a_p(t)$. The peripheral $a_p(t)$ functions combine in some fashion to produce a single *central* $a_\theta(t)$ *function*, from which information is extracted by the acquisition-rate model. The sensory threshold, Θ, is assessed before the combination of the peripheral $a_p(t)$ curves, subsequent to the combination, or in both locations.

Empirical Questions and Conclusions

The linear-filter model, in conjunction with the preceding assumptions, can be used to address a number of questions. The first is the location of the presumed sensory threshold Θ. Based on the empirical data, we concluded that two thresholds are required, one before and one after the combination mechanism.

A second issue addressed by our theory is the method by which the peripheral $a_p(t)$ curves combine. Candidate mechanisms include linear combination, in which the two $a_p(t)$ functions are simply added at each point in time, or quadratic summation, in which the two functions combine by the formula, $a_\theta(t) = [a_{p,l}(t)^2 + a_{p,r}(t)^2]^{1/2}$, where $a_{p,l}(t)$ and $a_{p,r}(t)$ are the sensory response functions from the left and right eyes, respectively. On the basis of evaluation of the model predictions, neither mechanism was clearly superior. However, on the basis of a number of considerations, we conclude that a variant of the quadratic summation model best characterizes the data.

A Final Puzzle

We conclude with a brief discussion of what we consider to be the weakest component of our theory. Recall that the information-extraction rate, $r(t)$, is assumed to be proportional to $a_\theta(t)$, the magnitude by which $a(t)$ exceeds threshold (see Equation 3). Although one could easily design a physical analog to such an assumption,[10] there is no a priori reason (that we know of) to expect it to be what it purports to be here, namely, an accurate description of the interface between the visual and cognitive systems. One would feel more comfortable with a weaker assumption, for example, that $r(t)$ is simply monotonically related to $a_\theta(t)$. However, no assumption except the strong one of proportionality implies our theoretical cornerstone: that memory performance is determined by $A_\theta(\infty)$.[11]

Yet, the theory describes the data extremely well, not only these present experiments but also for other experiments described by Loftus and Ruthruff (1994, involving stimulus intensity–stimulus duration trade-offs) and by Loftus et al. (1993, involving masked stimuli). This means that one of two things is true: (a) the proportionality assumption has some basis in neurological reality or (b) that neurological reality involves some other set of principles that imply a close approximation to

proportionality. Both of these possibilities warrant investigation and, indeed, the latter is currently undergoing rather extensive scrutiny in our laboratory.

---

[10] Imagine, for example, an urn filled with a gas whose density varies over time by some function a, such that at time t, density is $a(t)$ molecules/m³. Every c s, some constant volume, $V(V$ in m³), of gas from the urn is captured and placed somewhere else. The rate $r(t)$ of capturing gas molecules would then be $r(t) = (V/c)a(t)$ molecules/s; that is, rate of molecule extraction, $r(t)$ would be proportional to $a(t)$.

[11] The phrase *determined by* warrants some explanation. In our theory, as we have seen, performance, P, is proportional to above-threshold area. Such proportionality depends on the information-acquisition rate, $r(t)$, being proportional to a specific function of acquired information: $[1.0 - I(t)]$. Elsewhere (e.g., Loftus et al., 1993) we have shown that even when the assumed relation between $r(t)$ and $I(t)$ is weakened substantially, performance is still monotonically related to above-threshold area.

References

Ames, C., & Palmer, J. (1992). *Macintosh Experimental Software.* Unpublished manuscript.

Blake, R., & Fox, R. (1973). The psychophysical inquiry into binocular summation. *Perception and Psychophysics, 14,* 161–185.

Busey, T. A. (1992). Tachistoscopic software for Macintosh computers. *Behavior Research, Methods, Instruments, and Computers, 24,* 426–430.

Carlson, C. R., & Klopfenstein, R. W. (1985). Spatial-frequency model for hyperacuity. *Journal of the Optical Society of America: A, 2,* 1747–1751.

Di Lollo, V., & Dixon, P. (1992). Is the icon's worth apples and oranges? Some fruitful thoughts on Loftus, Duncan, and Gehrig (1992). *Journal of Experimental Psychology: Human Perception and Performance, 18,* 550–555.

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

Foley, J. M., & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research, 21,* 1041–1053.

Gottesman, J., Rubin, G. S., & Legge, G. E. (1981). A power law for perceived contrast in human vision. *Vision Research, 21,* 791–799.

Graham, N. (1977). Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. *Vision Research, 17,* 637–652.

Groner, M. T., Bischof, W. F., & Di Lollo, V. (1988). A model of visible persistence and temporal integration. *Spatial Vision, 3,* 293–304.

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

Legge, G. (1980). A power law for contrast discrimination. *Vision Research, 21,* 457–467.

Legge, G. E., Kersten, D., & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America: A, 4,* 391–404.

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology, 3,* 525–551.

Loftus, G. R., & Busey, T. A. (1992). Multidimensional models and iconic decay: Reply to Di Lollo and Dixon. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 556–561.

Loftus, G. R., Busey, T. A., & Senders, J. (1993). Providing a sensory basis for models of visual information acquisition. *Perception and Psychophysics, 54,* 535–554.

Loftus, G. R., Duncan, J., & Gehrig, P. (1992). On the time course of

perceptual information that results from a brief visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 530–549.

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). Extraction of information from complex visual stimuli: Memory performance and phenomenological appearance. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 139–191). San Diego, CA: Academic Press.

Loftus, G. R., & Ruthruff, E. (1994). A theory of visual information acquisition and visual memory with special application to intensity–duration trade-offs. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 33–49.

Massaro, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychological Review, 77,* 557–567.

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039–1042.

Pelli, D. G. (1981). *Uncertainty in visual detection and identification.* Unpublished doctoral dissertation, Cambridge University. Cambridge, England.

Roufs, J. A. J., & Blommaert, F. J. J. (1981). Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. *Vision Research, 21,* 1203–1221.

Rumelhart, D. E. (1970). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology, 7,* 191–218.

Sachs, M. B., Nachmias, J., & Robson, J. G. (1971). Spatial-frequency channels in human vision. *Journal of the Optical Society of America, 61,* 1176–1187.

Schurman, D. L., Eriksen, C. W., & Rohrbaugh, J. (1968). Masking

phenomena and time-intensity reciprocity for form. *Journal of Experimental Psychology, 78,* 310–317.

Shibuya, H., & Bundesen, 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. (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. (1979). Critical duration, supersummation, and the narrow domain of strength-duration experiments. *Behavioral and Brain Sciences, 2,* 279.

Sperling, G., Budiansky, J., Spivak, J., & Johnson, M. C. (1971). Extremely rapid visual search: The maximum rate of scanning letters for the presence of a numeral. *Science, 174,* 307–311.

Sperling, G., & Melchner, M. J. (1978). Visual search, visual attention, and the attention operating characteristics. In J. Requin (Ed.), *Attention and performance* (Vol. 7, pp. 675–686). Hillsdale, NJ: Erlbaum.

Sperling, G., & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America, 58,* 1133–1145.

Stoddard, P. K., & Loftus, G. R. (1988). An IBM XT-compatible, computer-based, slide-projector laboratory. *Behavior Research Methods, Instruments, & 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.

Watson, A. B. (1979). Probability summation over time. *Vision Research, 19,* 515–522.

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.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics, 33,* 113–120.

Appendix A

Visual Information Acquisition

Table A1
Glossary of Notations

Symbol	Definition	Unit	Depends on
Physical attributes			
d	Stimulus duration	ms	
t	Time since stimulus onset	ms	
ϕ	Stimulus contrast	percent (contrast)	
Performance measures			
p	Proportion digits recalled (corrected for 10% guessing probability)	proportion	
P	Transformed performance: $-\ln(1.0 - p/Y)$	ln(proportion)	
Regression parameters			
c_r	Estimated slope inverse	ms	
L	Estimated duration—axis intercept	ms	
Y	Estimated performance asymptote	proportion	
Linear systems parameters			
τ	Decay constant of the impulse response function	ms	
n	Number of stages in the impulse response function	integer	
Θ	Sensory threshold	percent (contrast)	
Linear systems functions			
$f(t)$	Stimulus contrast function	percent (contrast)	
$F(t)$	Integral from 0 to t of $f(t)$	ms	
$F(\infty)$	Total area under $f(t)$	ms	
$g(t)$	Impulse response function	percent (contrast)	τ, n
$G(t)$	Integral from 0 to t of $g(t)$	ms	τ, n
$G(\infty)$	Total area under $g(t) = 1.0$	ms	τ, n
$a(t)$	Sensory response function	percent (contrast)	$f(t), \tau, n$
$A(t)$	Integral from 0 to t of $a(t)$	ms	$f(t), \tau, n$
$A(\infty)$	Total area under $a(t)$	ms	$f(t), \tau, n$
$a_\Theta(t) = a(t) - \Theta$	Above-threshold sensory response function	percent (contrast)	$f(t), \tau, n, \Theta$
$A_\Theta(t)$	Integral from 0 to t of $a_\Theta(t)$	ms	$f(t), \tau, n, \Theta$
$A_\Theta(\infty)$	Total area under $a_\Theta(t)$	ms	$f(t), \tau, n, \Theta$
Extraction-rate parameter			
c_s	Reciprocal of raw sampling rate (also, reciprocal of proportionality constant between $A_\Theta(\infty)$ and P)	ms	
Extraction-rate and performance functions			
$r(t)$	Information-extraction rate function	ms^{-1}	$f(t), \tau, n, \Theta, c_s$
$I(t)$	Extracted-information function	proportion	$f(t), \tau, n, \Theta, c_s$
$I(\infty)$	Total extracted information	proportion	$f(t), \tau, n, \Theta, c_s$

Appendix B

Proof That the Independent-Sampling Model Implies Equation 8

Let the raw sampling rate be $a_\Theta(t)(1/c_s)$: that is, features are always sampled at a rate of $1/c_s$ per unit time modulated by the degree to which $a(t)$ exceeds threshold. Whenever a feature is sampled, it is transferred to short-term memory. At time t following stimulus onset, some proportion, $I(t)$, of the features have been sampled and are in short-term memory.

The remaining proportion, $[1.0 - I(t)]$, are new features, that is, features that have never been sampled. Accordingly, at time t following stimulus onset, the proportion of the $(1/c_s)$ sampled features that are new is $[1.0 - I(t)]$, which means that the rate of sampling new features, $r(t)$, is as follows:

$$r(t) = a_\Theta(t) \frac{[1.0 - I(t)]}{c_s}. \tag{B1}$$

Because $r(t)$ is the derivative of new features, $I(t)$ with respect to time, Equation B1 may be rewritten as,

$$\frac{d[I(t)]}{dt} = a_\Theta(t) \frac{[1.0 - I(t)]}{c_s}.$$

Rearranging terms,

$$\frac{d[I(t)]}{[1.0 - I(t)]} = a_\Theta(t) \frac{dt}{c_s}. \tag{B2}$$

Integrating both sides of Equation B2,

$$-\ln[1.0 - I(t)] = A_\Theta(t) \frac{t}{c_s} + k, \tag{B3}$$

where k is the constant of integration. When $t = 0$, $I(t) = 0$, which implies that $k = 0$. The total information acquired from the stimulus is obtained by setting t to ∞. Setting t to ∞, setting k to 0, and exponentiating Equation B3 yields,

$$1.0 - I(\infty) = e^{-A_\Theta(\infty)/c_s},$$

or, rearranging terms,

$$I(\infty) = 1.0 - e^{-A_\Theta(\infty)/c_s},$$

which is the text equation with $t = \infty$.

Appendix C

Details of Parameter Estimation

In general, parameter fitting was done by using a gridsearch procedure that computed the root-mean-square error (RMSE) for the fit of a given parameter set. A systematic, full-factorial search of all parameters provided the parameter set yielding the lowest RMSE. Table 1 provides a summary of the best fitting model parameters for all six experiments.

Experiments 1, 2, and 3

When fitting the Experiment 2 data, we assumed that each of the two presentations was processed independently, in other words, that the $a(t)$ curve for one presentation did not interact with the $a(t)$ curve for the second presentation. The use of a 250-ms interstimulus interval (ISI) justifies this assumption because a 50-ms presentation's $a(t)$ curve is essentially zero 250 ms after stimulus offset. This assumption was not made for Experiment 3 data.

Experiment 4

The gap-monoptic model fit of Experiment 4 required the same independent observations assumption as in Experiment 2. Again, this assumption was justified by the long ISI.

Experiment 5

To fit the Experiment 5 data, we explicitly assumed the cathode ray tube model, necessary because the stimuli were ramped on and off.

Experiment 6

To fit the Experiment 6 data, we assumed that the detection and identification tasks extracted information at different rates. Thus, we modeled the detection task with a Detect c_s and the identification task with an Ident c_s.

Appendix D

Theoretical Implications of Using a Computer Monitor as a Display Device
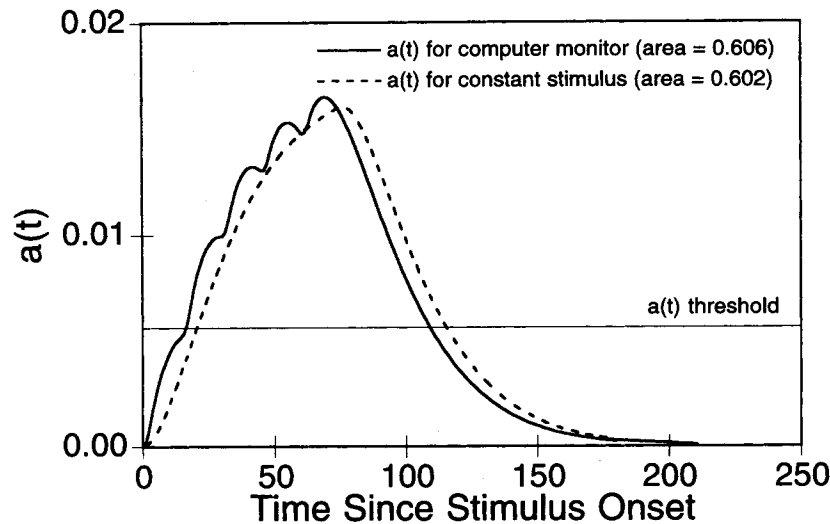

Figure D1. Sensory response curves generated by a constant stimulus (e.g., a slide projector) and a flickering stimulus (e.g., a cathode ray tube). Despite differences in the overall shapes of the two curves, the area differences are quite small.

At the heart of our theory is a *temporal low-pass filter,* which transmits low, but not high, temporal frequencies. The system's response to these high frequencies is indistinguishable from its response to constant stimuli (such as a slide projector).

A computer display monitor changes the image at 67 Hz, which, as we demonstrate here, produces temporal frequencies that are translated by the visual system into phenomenologically stable images. The exact shape of the sensory response function $a(t)$ depends somewhat on the value of the model parameter τ: For large τs (>30), the sensory response functions to a computer display and a projector display are virtually identical. As τ decreases to around 10, the response to a computer display begins to "scallop" somewhat (see Figure D1).

Although there appears to be a dramatic difference, the two curves share the same overall area and have nearly identical above-threshold areas. The only difference between the two curves arises where each curve crosses the sensory threshold. This difference appears slight and should not affect the model's predictions. To check the implications precisely, the combined Experiment 3 data were fit by a version of the model that incorporated the computer display assumptions.[12] The results from this analysis (see Table D1) show that although the model parameters changed slightly due in part to trade-offs between the parameters, there was no real change in the root-mean-square error. Thus, we conclude that from the theory's perspective, the visual system's response to a computer display may be treated as essentially identical to its response to a slide projector.

Table D1

Best Fitting Parameters for the Combined Data for Experiment 3

Display	τ	c_s	Θ	RMSE
Slide projector	13.9	0.653	2.192	0.035
Computer display	14.3	0.610	2.251	0.037

Note. RMSE = root-mean-square error. This table demonstrates the similarity of the linear filter's response to a slide projector display and a computer display.

[12] We assumed that the computer display generated a spike 3 ms in duration, scaled appropriately in contrast. Thus, for the first 3 ms, the $f(t)$ function had a contrast of 0.1871 and then had a contrast of 0 for the remaining 12 ms of the display. Note that this is an extreme assumption: Most points on the screen are not updated until the middle of the 15-ms interval, and altering the model to account for this would make computer display $a(t)$ functions more like projector display $a(t)$ functions.