A front end to a theory of picture recognition

GEOFFREY R. LOFTUS and JENNIFER E. McLEAN
University of Washington, Seattle, Washington

We describe the results of four picture-recognition memory experiments in which we systematically manipulated four variables: stimulus duration, stimulus contrast, the duration of a blank gap between successive presentations of the same stimulus, and the presence or absence of a noise mask that immediately followed stimulus offset. The patterns of obtained data confirmed a simple extension of a theory previously used to account for digit recall data. This theory consists of a low-pass linear-filter front end that generates a sensory response from the physical stimulus, followed by an information-sampling process whose instantaneous sampling rate is based in part on the sensory response magnitude. The data confirm both qualitative and quantitative theoretical predictions, some of which were previously untestable in digit recall tasks because of ceiling effects that were not present in our picture-recognition tasks. We describe the role of our theory within the broader family of picture-memory theories, and we briefly discuss our theory's unification of two salient facets of visual behavior: information acquisition on the one hand, and phenomenological appearance on the other hand.

At its most basic level, the standard information-processing view of visual perception and memory is the following. A visual stimulus contains information. When a person views a visual stimulus, information contained in the stimulus is encoded via a set of sensory, perceptual, and cognitive processes, resulting in a memory representation of that stimulus that serves as the basis for a response in a later memory task.

The kinds of mental processes that intervene between stimulus presentation and the memory test, along with the nature of the memory representation, are presumed to depend on the type of stimulus and task. The possible stimulus–task combinations range in complexity from the relatively simple (e.g., detection of a patch of light on a black background) to the relatively complex (e.g., delayed recognition of a naturalistic picture). Each stimulus–task combination presumably requires its own set of specialized mental processes. Even the simplest stimulus–task combination—detection of a patch of light—involves processes that have been examined in great detail, yet are still not completely understood. Picture recognition is much more complex because there are an unlimited number and variety of pictures in the visual world; thus, high-level mechanisms for pattern recognition and categorization that are not needed to detect a patch of light are almost certainly required.

Despite these differences in the kinds and complexity of processing involved in different stimulus–task combinations, there are also important similarities. In particular, regardless of the stimulus or task, the early stages of the visual system are generally assumed to produce a sensory representation of the stimulus that is closely linked to the physical properties of the stimulus. Then the information from the stimulus contained in the sensory representation must be acquired from it to be further processed and represented in memory in a way that is not necessarily as closely linked to the physical properties of the stimulus.

In this article, we focus on this early representation and on consequences of this early representation that do not depend on the task or the type of stimulus. By this, we mean the following, which we offer as a working hypothesis. All else being equal, eventual performance on any kind of visual memory task is limited by the nature of the early representation: If this representation is incomplete, then subsequent performance is imperfect, and manipulations of the nature of the early representation will affect performance. Thus, if one can devise a theory that is specific enough to describe the nature of the early representation in detail, then, all else held equal, the theory provides concomitantly specific predictions about any stimulus set being tested by any task.

In this article, we apply just such a theory. In particular, we extend to a complex task—delayed picture recognition—a theory that has dual roots in the vision literature and in the information-processing literature. The theory has been used extensively to account for performance in a much simpler task—immediate digit recall performance—across various stimulus conditions. It is our goal here to show how this theory can predict effects that have their origins in the precisely predictable nature of the initial representation, but which then extend through numerous encoding and decision processes, eventuating in precisely predictable effects on picture recognition performance.

Using the theoretical machinery of a linear filter, our theory focuses in detail on the front end of the perceptual/

Figure 1. Major theoretical components.

tasks. Because we confine our theoretical focus to sensory and information-acquisition stages, we report experiments wherein we manipulate variables such as stimulus duration and stimulus contrast that are likewise presumed to affect those stages. We show how certain nonobvious effects of these variables that are theoretically predicted and observed in a quite simple situation—immediate recall of briefly presented digits—are also theoretically predicted and are observed in a much more complex situation—delayed recognition of complex, naturalistic pictures. In addition, we test certain fundamental predictions of our theory that, for technical reasons that we describe below, cannot be tested in a simple digit recall paradigm, but that *can* be tested in a picture recognition paradigm.

The remainder of this article is organized as follows. First, we describe our theory and summarize how it has successfully predicted a number of findings. We then describe how the theory predicts several effects that have not been tested in the past, but that can be tested in a picture memory paradigm. We then provide a series of experiments in which these predictions are confirmed.

A THEORY OF VISUAL INFORMATION ACQUISITION

The theory to which we have been alluding is an extension of one put forth originally by Loftus and his colleagues (Busey & Loftus, 1994, 1998; Loftus, Busey, & Senders, 1993; Loftus, Duncan, & Gehrig, 1992; Loftus & Ruthruff, 1994). This theory provides an account of some of the general aspects of the visual processing that occurs between stimulus presentation and memory representation. It incorporates three principal components: first, a physical representation of the stimulus; second, a presumed automatic sensory response triggered by the stimulus; and third, transfer of visual information to semipermanent or permanent storage. We will sketch these components in turn.

Representation of the Physical Stimulus: $f(t)$

A stimulus is represented as a temporal function, denoted $f(t)$, that relates stimulus contrast to time (t) since stimulus onset. In a typical brief presentation, the representation might be as that shown in Figure 1A. Here a stimulus is presented for $d = 40$ msec at a constant $\Phi = 5\%$ contrast level.[2] We refer to $f(t)$ as the *stimulus temporal waveform*.

The Sensory Response: $a(t)$

The theory's next component is what we refer to as a *sensory response*. The sensory response may be conceptualized as summarizing the neural activity that is directly caused by the presence of the physical stimulus. Hence this sensory representation is linked closely to the physical characteristics of the stimulus: The magnitude of the sensory response is assumed to vary with t, the time since stimulus onset, and is hence a function, which we denote
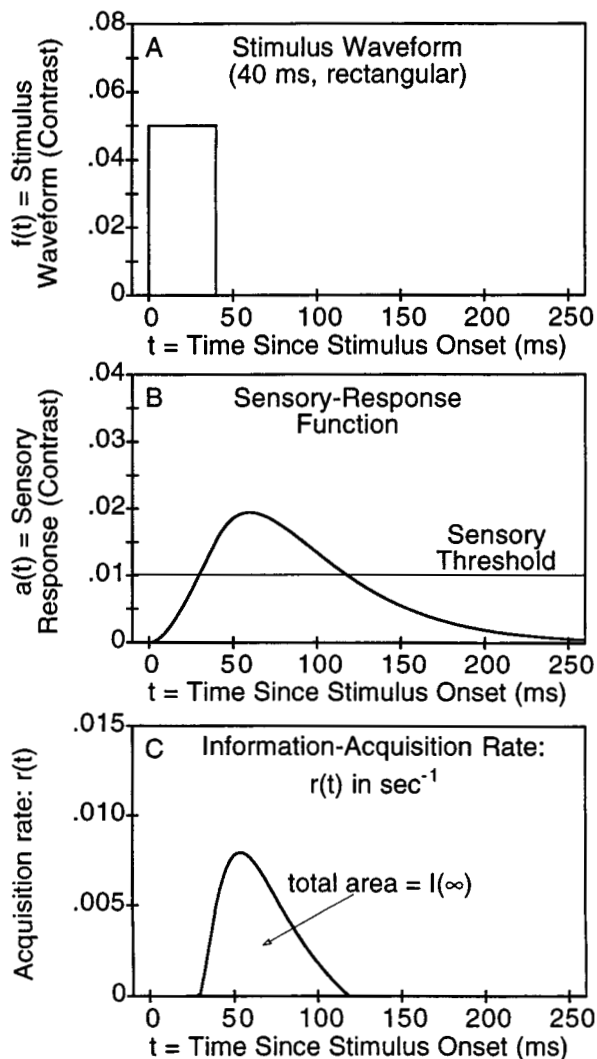
cognitive system; thus, the nature of the sensory response to a stimulus is well specified. As we move into cognition, our theoretical constructs become less well specified: we assume simply that information is extracted at an instantaneous rate based in part on the instantaneous sensory response magnitude, and that eventual performance is monotonically related to the resulting amount of acquired information. Object recognition and recognition decision processes are not specified, but rather are simply bundled together, along with everything else that must take place to get from original stimulus information to later memory task performance, into this monotonic function. A more complete model of picture perception and memory would, of course, include an explanation of these complex cognitive processes.[1]

We thus intend to demonstrate how a single theory can account for the early sensory and information-acquisition stages of visual perception for a wide range of stimuli and

$a(t)$. An example of this function—whose genesis we describe next—is shown in Figure 1B. (The horizontal line in Figure 1B labeled "Sensory Threshold" will be described below.)

The sensory response function is generated by assuming that the initial stages of the visual system act as a linear temporal filter that operates on $f(t)$, the stimulus temporal waveform. The mathematical properties of such a filter are described in detail elsewhere (e.g., Sperling, 1964a; Watson, 1986), as is the specific role of such a filter within the present theory (e.g., Busey & Loftus, 1994; Loftus & Ruthruff, 1994). The filter's operation can be described as follows.

A visual stimulus can, in principle, be presented arbitrarily quickly (e.g., the illumination provided by a lightening bolt lasts perhaps a microsecond). Like most physical systems, however, the visual system is incapable of reacting arbitrarily quickly to a stimulus. Instead the system's response lags behind and is temporally blurred in comparison with the stimulus, as can be seen by comparing the physical function in Figure 1A with its sensory counterpart in Figure 1B.

This kind of situation can be, and often is, modeled by a low-pass temporal filter. Such a filter efficiently passes the low temporal frequencies inherent in the physical stimulus, but passes higher temporal frequencies progressively less efficiently. The properties of a given filter can be characterized in either of two mathematically equivalent ways. In the *frequency* domain, one can specify the *transfer function*, which describes the degree to which the filter passes stimuli of varying temporal frequencies. In the *temporal* domain, one can specify the *impulse–response function*, which describes the system's response to an *impulse* (defined as a stimulus of infinite intensity, infinitesimal duration, and integral of 1.0). We denote the impulse–response function as $g(t)$. A commonly used filter in sensory theory treats the impulse response as a gamma probability density function. The reasons for using a gamma function in the present theory are provided elsewhere (e.g., Busey & Loftus, 1994). For present purposes, it is sufficient to note that a gamma function is a convolution of exponentials, defined by two parameters: n, a positive integer corresponding to the number of exponentials, and τ, a positive real number corresponding to the mean decay time of each exponential.

Linear Filters

A filter is *linear* if it has the property that its response to the superposition of two stimuli is equal to the sum of its responses to the two stimuli individually. Assuming a filter to be linear confers a major theoretical benefit: Knowing only the parameters of the filter (n and τ in this case) allows one to generate the sensory response function for a stimulus of any arbitrary temporal shape.

Given a linear filter defined by the impulse–response function, $g(t)$, the filtered response, $a(t)$, to the stimulus, $f(t)$, is the convolution of $f(t)$ with $g(t)$; that is,

$$a(t) = (f * g)(t) = \int_{-\infty}^{\infty} f(t-u)\,g(u)du, \qquad (1)$$

where "$*$" represents convolution.

Before we proceed, a few brief remarks about linearity are in order. A great deal of work in the sensory literature has revolved around the degree to which the visual system can be construed as linear and around the nature of numerous nonlinearities that are observed under various circumstances (see, e.g., Di Lollo & Bischof, 1995; Sperling, 1964a, 1979; Sperling & Sondhi, 1968; Watson, 1986 provides an extensive overview). It is generally accepted that (1) the system is linear and (2) the impulse–response function is monophasic, only when stimulus contrast and stimulus luminance are relatively low. How low is "low"? This is a matter of some question. However, Loftus and his colleagues, using a digit recall task, have obtained excellent fit to the predictions of a model which assumed linearity with contrasts up to about 20%—the higher range of contrast values used in the present experiments. But an assumption of linearity could not account for similar digit recall data reported by Shibuya and Bundesen (1988) wherein much higher contrast levels were used.

Sensory Threshold

The concept of a *sensory threshold*, which has to do with distinguishing informational signal from background noise, is well-established in the vision literature, and we incorporate that key idea into our theory as follows. We assume there to be a sensory threshold (which we designate Θ) such that no postsensory processing occurs except when the magnitude of the sensory response exceeds this threshold. The threshold is indicated in Figure 1B as a horizontal line. For notational convenience, we define a new function, $a_\Theta(t)$ as follows:

$$a_\Theta(t) = \begin{cases} 0, & a(t) < \Theta \\ a(t) - \Theta & a(t) \geq \Theta \end{cases}. \qquad (2)$$

Thus, if $a_\Theta(t)$ fails to exceed zero, no subsequent processing occurs.

Information Acquisition: $r(t)$

Based in part on the magnitude of the sensory response, or more specifically, $a_\Theta(t)$, information is assumed to be acquired from the stimulus and placed into working memory where it can be processed further. We assume that at any time, t, there is an instantaneous rate of such information acquisition, which we denote $r(t)$.

Conceptualization of $r(t)$

We define $I(t)$ to be the proportion of stimulus information acquired from the stimulus at time t. Therefore, $[1.0 - I(t)]$ is the proportion of yet-to-be-acquired information. With other theorists (e.g., Bundesen, 1990; Rumelhart, 1970; Shibuya & Bundesen, 1988; Townsend, 1981), we assume $r(t)$ (which, of course, is dI/dt, the derivative of $I(t)$, with respect to time) to be proportional to the product of two quantities: first, the magnitude of the above-

threshold sensory response, and second, the proportion of yet-to-be-acquired information; thus,

$$r(t) = a_\Theta(t)[1.0 - I(t)]/c, \quad (3)$$

where $1/c$ is the constant of proportionality. Hence, $r(t)$ is zero unless $a_\Theta(t)$ exceeds zero, and $I(t)$ can approach, but can never exceed, 1.0. A typical $r(t)$ function is shown in Figure 1C.

Note that $I(t)$ is, by definition, the integral over time of $r(t)$, the information-acquisition rate. Therefore, the proportion of stimulus information that has been acquired at any time, t, is

$$I(t) = \int_0^t r(t)dt. \quad (4)$$

Computation of $r(t)$

It is easily shown that, given Equations 3 and 4,

$$I(t) = 1.0 - e^{-A_\Theta(t)/c} \quad (5)$$

where

$$A_\Theta(t) = \int_0^t a_\Theta(t)dt. \quad (6)$$

If we combine Equations 3 and 5, therefore,

$$r(t) = \left[a_\Theta(t)e^{-A_\Theta(t)/c}\right]/c. \quad (7)$$

Total Acquired Information: $I(\infty)$

Of interest is the *total* proportion of information acquired from the stimulus on a given trial. This quantity, denoted $I(\infty)$, may be obtained by letting t approach ∞ in Equation 5 and is, of course, the total area under the $r(t)$ function, as indicated in Figure 1C. Total acquired information is obtained by

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

Note here that $A_\Theta(\infty)$ is simply equal to the total above-threshold area under $a(t)$.

Performance

For the moment, we make a very weak assumption relating memory performance in some experimental condition to per-stimulus acquired information in that condition: that a monotonic function, m, relates the two. That is,

$$p = m[I(\infty)], \quad (9)$$

where m is monotonic. Below, we shall strengthen this assumption in order to allow a quantitative comparison between theoretical predictions and observed memory performance.

The Identity Property

A key property of a linear filter (assuming the impulse–response function to be nonnegative, as it is in this theory) is that the area under the filtered function, $a(t)$ in this case, is equal to the area under the original stimulus temporal input function, $f(t)$. This means that two stimuli with identical areas under $f(t)$ must engender corresponding $a(t)$ functions that, while not necessarily having the same shape, must nonetheless have the same total area. We refer to this property of our linear-filter model as the *identity property*.

Suppose that the threshold value were negligible (i.e., that $\Theta \approx 0$), which is a state that we achieve in the experiments described below. In that case, $a_\Theta(t) \approx a(t)$, and stimuli with equal areas under $f(t)$ would have approximately equal areas under $a_\Theta(t)$—that is, equal values of $A_\Theta(\infty)$. Such stimuli would, by Equations 8 and 9, yield approximately equal $I(\infty)$ values and, hence, approximately equal performance.

Theory: Summary

Let us briefly recapitulate the theory as we have described it thus far. Steps 1–5 below can be construed as a theoretical path from the observable input stimulus in a given experimental condition to the observable response probability for that condition.

1. The (observable) stimulus in some experimental condition is represented by $f(t)$, which relates stimulus contrast to t, the time since stimulus onset. The shape of $f(t)$ is set by the experimenter, and it defines a particular experimental condition.

2. The sensory response function, $a(t)$, is generated by applying a linear filter to $f(t)$. Specifying τ and n, the filter parameters, allows computation of $a(t)$ from $f(t)$, via Equation 1.

3. Specifying Θ, the value of the sensory threshold, in addition to $a(t)$ from Step 2, allows computation of $a_\Theta(t)$, the above-threshold magnitude of $a(t)$, by Equation 2, and $A_\Theta(\infty)$, the total area under $a_\Theta(t)$.

4. $I(t)$ is defined to be the amount of information acquired at time t, and the information-acquisition rate, $r(t) = dI/dt$, is proportional to the product of $a_\Theta(t)$ and $[1.0 - I(t)]$. Specifying the value of the proportionality constant, c, along with $A_\Theta(\infty)$ from Step 3, allows computation of $I(\infty)$, the total amount of acquired information per trial, via Equation 8.

5. Observable performance, p, is monotonically related to $I(\infty)$ as per Equation 9. Computing numerical predicted performance values requires, of course, specification of the function m.

Theoretical Predictions

As indicated, the theory has been applied to a number of digit recall paradigms. The typical task has been a simple one: a string of four low-contrast (approximately 3% contrast) digits is briefly presented to an observer who then attempts to recall them in correct order, guessing if necessary. The dependent variable is percent correct digits adjusted for the 10% guessing rate. Three effects of particular relevance to the present work are termed the *liftoff* effect, the *gap* effect, and the *contrast* effect. We will now
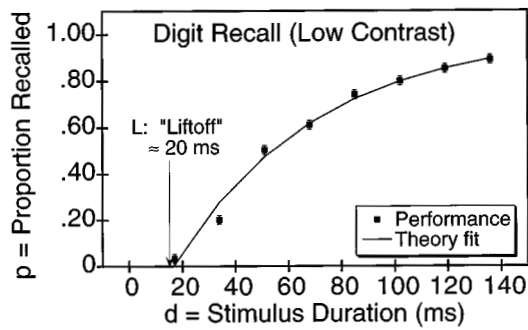
Figure 2. Digit recall: The liftoff effect. From "Sensory and Cognitive Components of Visual Information Acquisition," by T. A. Busey and G. R. Loftus, 1994, *Psychological Review*, *101*, p. 459. Copyright 1994 by the American Psychological Association. Adapted with permission.

describe these effects in turn, along with our theoretical accounts of them.

The Liftoff Effect

Figure 2 (adapted from Busey & Loftus, 1994) shows the result of a digit recall experiment in which stimulus duration was varied from approximately 10 to 140 msec; here performance is plotted as a function of stimulus duration. The symbols represent the observed data (with standard errors). The solid line represents the theory's predictions, in this and the next two data figures in this section.[3] The *liftoff effect* refers to the finding that there is some critical stimulus duration (in this instance, approximately 20 msec) below which performance is at zero (hence, "liftoff" is the duration at which the curve lifts off from the floor). This effect is quite pervasive, and others (e.g., Rumelhart, 1970; Shibuya & Bundesen, 1988; Townsend, 1981) have included it as a "delay parameter" in models of the task.

The theory's account of the liftoff effect is illustrated in Figure 3, which shows sensory response functions resulting from five exposure durations that range from 10 to 160 msec. The sensory threshold, Θ, is also shown. Note that the two sensory response functions corresponding to the two shortest durations (10 and 20 msec) do not exceed threshold. Because information acquisition occurs only when $a(t)$ exceeds threshold (see Equations 2–3), no stimulus information is acquired during these short-duration conditions, and memory performance for them is accordingly predicted to be zero. As duration increases beyond 20 msec, above-threshold area exceeds zero and increases; hence performance likewise increases.

The Gap Effect

Figure 4 (adapted from Busey & Loftus, 1994) shows the result of a digit recall experiment in which total stimulus exposure duration is constant (75 msec), but in which the temporal configuration of the presentation varied. In particular on a given trial, the digits were pre-

sented initially for 30 msec, followed by a variable-duration blank gap, followed by the same digits again for 45 msec. Gap duration varied from 0 to approximately 100 msec. It is apparent that performance declines with increasing gap duration.

The theoretical predictions are represented by the solid line in Figure 4. The theory's explanation of the gap effect is illustrated in Figure 5, the three panels of which show sensory response functions for three gap durations. The physical stimuli—the $f(t)$ functions—are depicted as dashed rectangles.

To understand the predictions embodied in Figure 5, let us momentarily ignore the sensory threshold. As we have noted, a salient property of a linear filter (that has a nonnegative impulse–response function) is that physically different stimuli which have identical areas under their $f(t)$ functions (e.g., the three stimuli in the Figure 5 example) will likewise have identical areas under their filtered (sensory response) functions. Thus, the three sensory response functions in the three Figure 5 panels, although having quite different shapes, all have identical total areas under the curves. Accordingly, if there were no threshold, performance would, by the identity property, be independent of gap duration.

However, the introduction of a threshold, and the associated dependence of performance on *above-threshold*, area changes this rule. In particular, when there is a threshold, "tall is good." By this we mean that, given two differently shaped curves with identical areas, the taller narrower curve provides more above-threshold area than does the shorter, wider curve, as is indicated in Figure 5. Because the greater the gap duration, the shorter and wider the resulting sensory response function, above-threshold area decreases with gap duration, thereby leading to lower performance.

The Contrast Effect

Figure 6 (adapted from Loftus & Ruthruff, 1994) shows the result of a digit recall experiment in which both stimulus contrast and stimulus duration were varied. Each of the four curves corresponds to a particular contrast and shows performance as a function of duration.
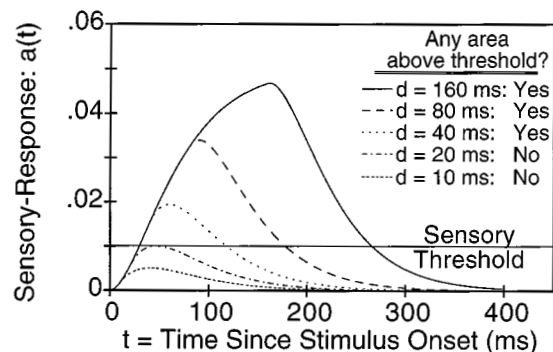

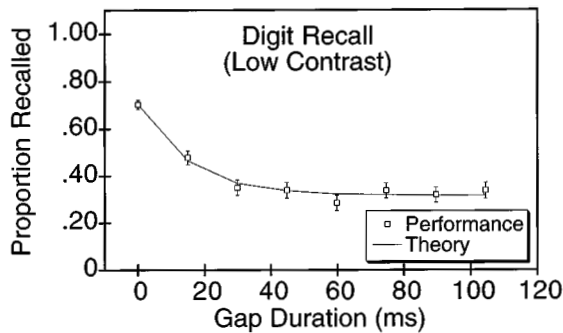**Figure 3. Digit recall: Liftoff effect explanation.**

Figure 4. Digit recall: The gap effect. From "Sensory and Cognitive Components of Visual Information Acquisition," by T. A. Busey and G. R. Loftus, 1994, *Psychological Review*, *101*, p. 459. Copyright 1994 by the American Psychological Association. Adapted with permission.

Again, solid lines show theoretical predictions. Two effects of decreasing contrast are apparent: First, the liftoff duration increases, and second, performance increases more slowly with increasing duration (i.e., contrast and duration act multiplicatively).

The theoretical account of both these effects is quite straightforward. By Equation 1, stimulus contrast has the effect of *scaling* the sensory response function. That is, increasing contrast by some factor, k, corresponds to multiplying the height of the stimulus function, $f(t)$, and hence $a(t)$, the convolution of $f(t)$ and $g(t)$, by the same factor. This means that as contrast decreases, duration must increase in order for the sensory response function to reach the sensory threshold (cf. Figure 3)—hence the theory's account of the inverse contrast–liftoff relation. As duration is increased by a given amount, the corresponding increase in the area under the sensory response curve is smaller, the lower the contrast—hence the theory's account of the multiplicative relation between duration and contrast.

Threshold-Driven Effects

Two effects that we have just described—the liftoff effect and the gap effect—result directly from the putative sensory threshold. By including such a threshold, the theory is capable of accounting for the effects quite well, as is demonstrated by the theoretical predictions shown in Figures 2, 4, and 6.

Threshold-Less Predictions

Suppose, however, that there were no threshold. In such a case, the theory would make a variety of strong qualitative predictions. Two such predictions that are relevant to the present experiments are the following:

1. Liftoff would be zero; that is, any curve relating memory performance to stimulus duration would pass through the origin.

2. There would be no gap effect. As discussed above, a consequence of a linear filter that maps $f(t)$ to $a(t)$ is the identity property: To reiterate, any two equal-area $f(t)$

functions (e.g., two stimuli differing only in gap duration) would yield corresponding equal-area sensory response functions (see Figure 5) and hence equal performance.

Confirmation of such threshold-less predictions would be valuable for three reasons.

First, the predictions are quantitatively precise. Liftoff is predicted to be exactly zero, and the gap effect prediction is predicted to be exactly null.

Second, the predictions are independent of the theoretical and experimental parameters. By this, we mean the following. If a threshold is assumed, the exact nature of both the gap and the liftoff effect depend on the form of the impulse–response function, the magnitude of the threshold, the rate of information acquisition, the contrast and duration of the stimulus, and the nature of the monotonic function relating performance to $I(\infty)$. If, on the other hand, there were no threshold, such effects would, as indicated above, not depend on anything and, accordingly the observation of, or failure to observe, the effects would constitute, respectively, strong confirmation, or strong disconfirmation, of the model.

Finally, it is worth pointing out that confirmation of the null gap effect prediction is counterintuitive in the following sense. It is easy for an observer in these experiments to distinguish between, say, a zero-gap condition and a 100-msec gap condition. One would predict, on a strictly
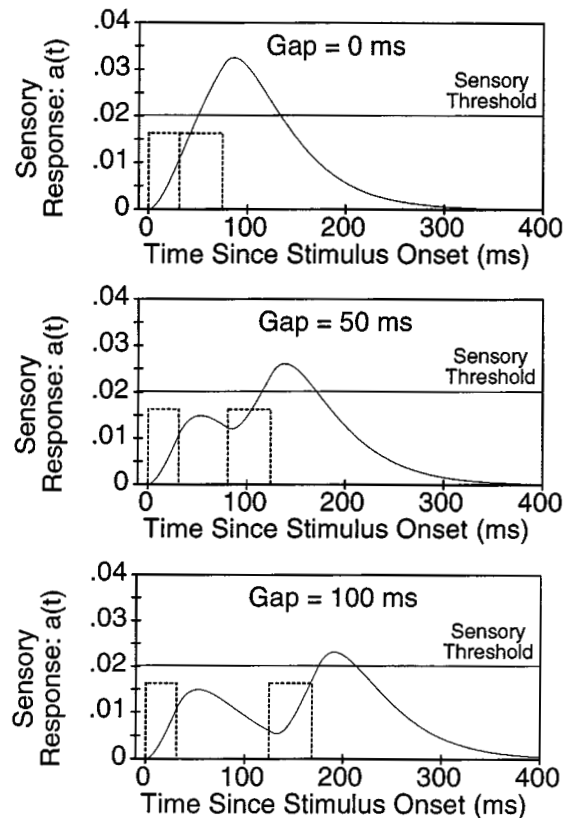

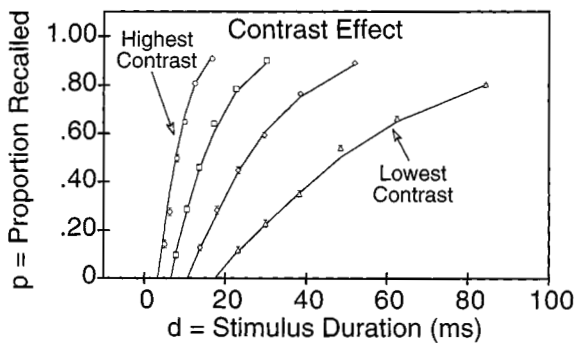**Figure 5. Digit recall: Gap-effect explanation.**

Figure 6. Digit recall: The contrast effect. Symbols on curve indicate data; solid lines indicate theory fit. From "A Theory of Visual Information Acquisition and Visual Memory with Special Application to Intensity-Duration Tradeoffs," by G. R. Loftus and E. R. Ruthruff, 1994, *Journal of Experimental Psychology: Human Perception & Performance*, **20, p. 37. Copyright 1994 by the American Psychological Association. Adapted with permission.**

intuitive basis, that two conditions that are very different phenomenologically would also be different with respect to eventual memory performance—much as, for instance, if one entered a strange country, one would expect that two coins of different appearance would probably be worth different amounts of money.

Rendering Threshold Irrelevant

So if there were no threshold, worthwhile experimental predictions would ensue. But how could we *test* such predictions? Obviously, if a sensory threshold is an intrinsic component of the perceptual system, one cannot experimentally remove it. However, it is possible to substantially diminish any threshold-driven effect by increasing stimulus contrast. Figure 7 illustrates why this is so. Here sensory response functions are shown for two stimuli: one of very low contrast (typical of the digit stimuli that we have been discussing), and the other, of higher contrast. With the low-contrast stimulus, much of the sensory-response function is below threshold, which would cause threshold-driven effects of high magnitude. With the high-contrast stimulus, however, there is proportionally less area below threshold, which would substantially diminish the magnitude of threshold-driven effects.

These considerations suggest an obvious set of experiments in which digit recall performance is measured following high-contrast stimulus presentation. However, it is not practical to carry out such experiments because such a contrast increase would lead to uninterpretable performance ceiling effects. It is partly for this reason that we carried out the present experiments using a picture recognition task in place of a digit recall task. Past pilot work indicated that even if pictures were presented at study at high contrast levels, eventual recognition performance would still be substantially below ceiling.

Extension of the Theory to Picture Recognition

Although the theory that we have described has been shown to account quite well for digit recall data, it is not tied to this particular stimulus and task. In the theory, we have characterized the stimulus simply in terms of its basic temporal characteristic: variation of contrast over time. This characterization applies equally well to other visual stimuli, be they simple, like luminance disks, or complex, like naturalistic pictures. However, using a picture recognition task to test the theory's predictions requires some auxiliary assumptions to cope with the added complexity.

Simplified Assumptions About
Picture Perception and Picture Recognition

As we have already discussed, picture perception and picture memory entail complex mental processes that have been the focus of a great deal of theoretical work (e.g., Biederman, 1993; Hummel & Biederman, 1992; Marr, 1982; Snodgrass & Feenan, 1990). Recognition memory (of pictures or anything else) is similarly complex (e.g., Eich, 1982; Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1993; Raaijmakers & Shiffrin, 1992). In extending our theory to picture recognition— that is, to the intersection of these two sets of complex processes—we simplify vastly via the following auxiliary assumptions.

1. Pictures are assumed to be identical to the already described digit stimuli (or to any other visual stimulus) in the following respects. A static picture is physically represented by a stimulus-contrast temporal waveform that is linearly filtered to produce a sensory response function with an associated sensory threshold. Information is acquired from the stimulus at a rate that is proportional to the product of above-threshold sensory response magnitude and proportion unacquired information. On any trial, the total proportion of acquired information about the picture shown on that trial can be computed and is denoted $I(\infty)$. This acquired information results from *per-*
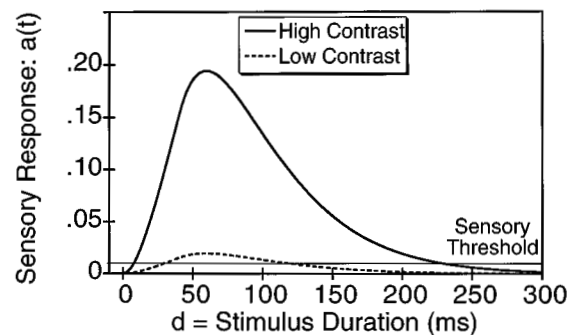

Figure 7. An illustration of substantially raising contrast: Any threshold-driven effects are less for the high-contrast than for the low-contrast stimulus.

ceptual processing (as defined by Intraub, 1980, 1984; Loftus, Hanna, & Lester, 1988; and Potter, 1976).

2. Acquired information is assumed to then serve as the basis for subsequent *conceptual processing* (again as defined by Intraub, 1980, 1984; Loftus et al., 1988; and Potter, 1976). The more information acquired from the picture, the greater the degree of conceptual processing accorded the picture.

3. The degree of conceptual processing is assumed to be monotonically related to recognition memory performance. This implies that $I(\infty)$ and memory performance are monotonically related, as has been previously assumed. We emphasize a critical consequence of monotonicity—the identity property described above—which is that, when threshold effects are negligible, two different physical stimuli having the same area under their $f(t)$ functions are predicted to have equal performance.[4]

4. Finally, we assume that within the bounds of the present experiments (wherein, for example, only a single eye fixation is allowed on a picture shown at study), performance asymptotes at some less-than-perfect value that is determined by limitations of postperceptual and retrieval processes. The asymptote's value, denoted Y, is a free parameter.

Risking redundancy, we would like to forestall confusion by reiterating one aspect of our theoretical extension to picture memory. In focusing on effects of sensory variables—particularly stimulus temporal waveform—we have essentially ignored numerous well-known effects of stimulus complexity, stimulus organization, and top-down processing effects (e.g., effects of priming and attention). By ignoring these effects, we do *not* mean to suggest that they do not exist or that they are unimportant for picture perception and picture memory. We mean, simply, that if these other effects are held constant, the sensory variables will be observed to influence picture memory, in ways that are described precisely by our theory.

Quantitative Predictions

As described so far, our theory is incapable of predicting exact performance in a picture recognition experiment, because the function relating $I(\infty)$ to performance is assumed only to be monotonic, which involves an unspecifiable number of parameters.

The experiments that we describe below are designed to test a number of what we refer to as "quasi-qualitative" predictions. By this, we mean the following. In our experiments, we compare recognition memory for low-contrast and high-contrast pictures. Above, we described "threshold-driven effects," and we argued that such effects diminish with high-contrast stimuli. If our low-contrast stimuli involved a threshold, while our high-contrast stimuli involved *no* threshold, then many of our predictions (e.g., of a decreasing gap effect for low-contrast pictures, but *no* gap effect for high-contrast pictures) would be *qualitative* in the sense that the predictions would hold for

all parameter values, stimulus properties, and monotonic functions assumed to relate $I(\infty)$ and performance. However, because we cannot argue that our high-contrast stimuli involve *no* threshold, the best we can do is predict that such stimuli will behave very much as if there were no threshold; hence the term "quasi-qualitative."

Because this situation is not entirely satisfying, we decided to strengthen the theory's monotonic assumption in order to provide an existence proof—to demonstrate that there is at least one quantitative instantiation of it that is capable of quantitatively describing our data. In the absence of a much more complete picture-perception and picture recognition theory, the nature of such strengthening is somewhat arbitrary. The first necessary ingredient is a single measure of memory performance that combines hit and false alarm probabilities. For reasons having do with the designs of our experiments (provided below), it is not especially important to our results or conclusions which measure we choose, as long as it is zero when memory is zero. An obvious candidate for such a measure is d'. However d' is frequently uncomputable when there are few data points (e.g., for a single subject in a single condition), which leads to occasional values of 1.0 for a hit probability and/or 0 for a false alarm probability. Thus we opted instead for the measure

$$p = \frac{p(\text{H}) - p(\text{FA})}{1.0 - p(\text{FA})}, \qquad (10)$$

where $p(\text{H})$ and $p(\text{FA})$ represent the observed hit and false alarm probabilities. To complete the quantitative version of the theory, we assume that $p = Y[I(\infty)]$, where, recall, Y is the performance asymptote.

Free Parameter Summary

The theory as we have described it thus far has five free parameters. The two linear-filter parameters, n and τ, generate the sensory response function, $a(t)$, from the observable temporal stimulus waveform, $f(t)$. The sensory threshold, Θ, allows calculation of above-threshold sensory response magnitude. The constant of proportionality, $1/c$, influences how fast task-relevant information is acquired from the stimulus (see Equation 2) and, in conjunction with n, τ, and Θ, allows calculation of $I(\infty)$ in a given condition. Finally, the performance asymptote, Y, allows prediction of performance from $I(\infty)$.

EXPERIMENTS

We report four experiments, each of which involved a sequence of four study–test blocks. In the study phase of each block, 24 target pictures—black-on-white line drawings depicting naturalistic scenes—were presented under varying conditions. In the test phase, which immediately followed each study phase, the 24 targets were randomly intermingled with 24 distractor pictures chosen from the same stimulus pool. The subjects responded

"old" or "new" to each test picture. The specifics of each experiment revolved around stimulus manipulations carried out during the study phases and were as follows.

In Experiment 1, we varied contrast and stimulus duration in order to examine the liftoff duration for high-contrast pictures and low-contrast pictures. The theory predicts that low-contrast pictures would, like the low-contrast digit stimuli, yield a liftoff value greater than zero, whereas high-contrast pictures would yield a liftoff value of zero or close to zero.

In Experiment 2, we held total stimulus duration constant (at 60 msec) and varied gap duration for high- and low-contrast pictures. The theory predicts that low-contrast pictures would, like the low-contrast digit stimuli, yield a declining function relating performance to gap duration, whereas high-contrast pictures would yield a gap-duration effect of zero or close to zero.

Experiment 3 was a control for the possibility that memory performance for the high-contrast pictures of Experiment 2 was at a maximum value, thereby producing the lack of high-contrast gap effect that we did in fact observe in Experiment 2. In Experiment 3, we replicated the high-contrast conditions of Experiment 2 both for the Experiment 2 total exposure duration (60 msec) and also for a longer total exposure duration (100 msec). We reasoned that if in Experiment 2 high-contrast performance was at a maximum, it could not be further increased by an exposure-duration increase; hence the finding of an exposure-duration effect would rule out this explanation.

In Experiment 4, we introduced a noise mask in order to test certain to-be-described theoretical predictions about masking effects.

Experiment 1
Variation of Stimulus Contrast and Stimulus Duration

The major goal in Experiment 1 was to investigate the liftoff value under conditions of low and high contrast. We did so by measuring recognition performance for pictures studied at six durations ranging from 5 to 160 msec. As indicated above, the most straightforward application of our theory to picture recognition implies that under low-contrast conditions, liftoff should be approximately the same as that observed in prior (low-contrast) digit recall experiments: approximately 20 msec. Under high-contrast conditions, however, the theory predicts that liftoff should be close to zero.

Method

Much of the methodology was the same over all four experiments. We will describe the Experiment 1 methodology in detail, emphasizing features common to all experiments. In each subsequent experiment, we will describe only the methodology unique to that experiment.

Subjects. In Experiment 1, 144 University of Washington undergraduates participated in a single 50-min session for extra course credit. They were run in 24 groups of 6 subjects per group.

Materials. Stimuli were 196 simple black-on-white line drawings of naturalistic scenes prepared as 35-mm slides and randomly placed in four trays of 48 slides per tray. The drawings were of common visual scenes such as a table set with a teapot, cups, fruit, and flowers; a man making pottery on a pottery wheel; a clown holding a birthday cake and balloons in front of two smiling children; and a woman holding a cat and looking through a rain-streaked window. The visual angle subtended by each picture ranged from 28.8° horizontally and 20.3° vertically to 19.4° horizontally and 13.7° vertically, depending on where a subject sat.

Contrasts were chosen as follows. Low contrast was chosen in such a way that the pictures were identifiable but barely so. High contrast was chosen so as to be as high as possible without entailing ceiling effects. Luminance of the black and the white regions of the pictures were measured over a sample of slides. In the high-contrast condition, mean luminance of the black regions was 9.0 cd/m², and of the white regions 14.9 cd/m², yielding a contrast of 23% [contrast measured as $(w - b)/(w + b)$]. In the intermediate (test phase) contrast condition, mean luminance was 7.1 cd/m² and 9.5 cd/m² for the black and the white regions, respectively, yielding a contrast of 14%. In the low-contrast condition, mean luminance was 6.1 cd/m² and 6.5 cd/m² for the black and the white regions, respectively, yielding a contrast of 3%. An ever present adapting field was measured to be 5.7 cd/m².

Apparatus. All stimuli were displayed via Kodak projectors equipped with Gerbrands tachistoscopic shutters. A random access projector was used to display the stimuli, while standard carousel projectors were used to present a constant, uniform adapting field and a fixation point that initiated each trial. Each projector intensity was controlled by attenuating projector luminance with a Wrattan neutral density filter. Each subject made test responses on an individual numeric keypad. All display equipment was enclosed in a soundproof box. All display and response collection was controlled by an IBM-compatible computer system described by Stoddard and Loftus (1988).

Design and Procedure. Each 6-subject group in Experiment 1 participated in a single experimental session. A session consisted of a study phase followed by a test phase for each of the four slide trays (thus each slide tray constituted an independent replication). Subjects were fully informed at the outset about the sequence, timing, and general nature of all events that would occur.

Specific procedures for each tray were as follows:

Study procedures. Within each stimulus tray, 24 target stimuli were presented during the study phase—1 target during each of 24 successive study trials. The stimulus on each trial was displayed in one of six conditions corresponding to six exposure durations: 5, 10, 20, 40, 80, or 160 msec. These conditions were presented in random order with the restriction that each condition occurred twice within each of two sequential 12-trial blocks constituting the 24-trial study phase. Prior to the beginning of a study phase, subjects were instructed to try to remember all of the upcoming target pictures.

Ideally, stimulus contrast, like stimulus duration, would have been varied randomly over study trials within each study phase. Equipment limitations precluded this, however, and required that contrast be varied across slide trays and be constant within each slide tray. This was done in a fashion that will be described in more detail below in conjunction with counterbalancing procedures.

The sequence of events on each study trial was as follows. First, a 500-msec, 1000-Hz tone signaled the subjects to fixate a dim spot that concurrently appeared at the center of the viewing field. The stimulus, superimposed over the adapting field, was then presented for one of the six durations. The stimulus onset asynchrony (SOA) from trial n to trial $n + 1$ was 3,000 msec.

Test procedures. The test phase for each tray immediately followed the study phase for that tray. In the test phase, all target stimuli from the just shown stimulus tray were randomly intermingled with an equal number of distractors, and the resulting test stimuli were shown one at a time. The target–distractor ordering was different for the different stimulus trays but, for each tray, was identi-
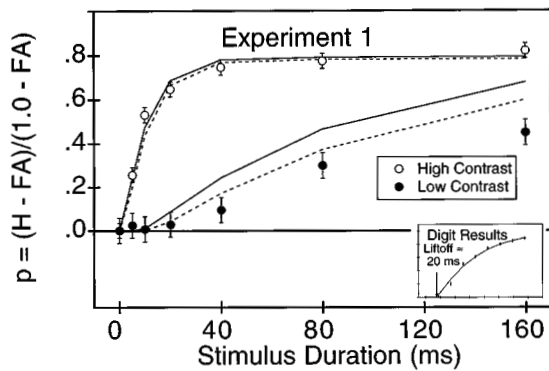
Figure 8. Data from Experiment 1: Top curves represent high-contrast data and bottom curves represent low-contrast data. Solid and dashed lines show theoretical fits as described in the text.

cal for all the groups throughout the experiment. Subjects judged each test stimulus to be old or new by pressing one of two designated keys on a response box. The subjects were cautioned that of the 48 test pictures that they would see, 24 had been presented during the study phase; hence they should respond "old" and "new" about equally often. Test picture contrast was set at approximately the mean of the two study contrasts

The sequence of events on each test trial was as follows. A test picture appeared, followed, after a 100-msec pause, by a 150-msec beep. This beep signaled the subjects that they could now respond. The test picture then remained visible until all subjects had responded. After all the subjects had responded, there was a 500-msec pause prior to the onset of the next test trial.

Counterbalancing. The stimuli were counterbalanced across target/distractor and across study conditions as follows. Within each slide tray, the 48 slides were randomly divided into two sets (Set A and Set B) of 24 slides per set. For 12 of the 24 groups (the odd numbered groups), Set A slides served as targets and Set B slides served as distractors, whereas the reverse was true for the remaining 12 (even numbered) groups. Thus each stimulus occurred as a target for half the experimental groups and as a distractor for the other half. Each stimulus occurred the same number of times in each of the 2 contrasts × 6 durations = 12 study conditions over the 12 groups in which it appeared as a target.

As noted, contrast level over the four trays was blocked by tray. For a given group, Trays 1–4 were presented at high, low, low, and high contrast levels, respectively (HLLH configuration), or low, high, high, and low, respectively (LHHL configuration). Half the even numbered groups and half the odd numbered groups received each of these two configurations.

Results

Mean hit probabilities (across the four trays and 144 subjects) were measured for each of the 12 exposure duration × contrast conditions. We note that, because all duration conditions were randomly intermingled within each tray at study, there were only two false alarm probabilities: one for the high-contrast condition, which was .200, and the other for the low-contrast condition, which was .295. Based on the logic embodied in Equation 10, a performance measure, p_{jk}, was computed for each duration × luminance condition, jk, using the formula

$$p_{jk} = \frac{p(\mathrm{H}_{jk}) - p(\mathrm{FA}_k)}{1 - p(\mathrm{FA}_k)}, \qquad (11)$$

where $p(\mathrm{H}_{jk})$ is the hit probability for condition jk, and $p(\mathrm{FA}_k)$ is the false alarm probability for contrast condition k.

Figure 8 shows the principal results of Experiment 1. Here, performance is shown as a function of exposure duration with different curves shown for high and low contrast. The error bars represent 95% confidence intervals computed as suggested by Loftus and Masson (1994). (Results subsequently reported in this article as $X \pm Y$ likewise denote mean ±95% confidence interval.) The solid and dashed lines represent best theoretical fits that will be described below. The small inserted graph in the lower right shows corresponding digit recall data (Busey & Loftus, 1994). We chose to show data for distractor pictures and data for target pictures on the same duration scale. In this sense distractors are simply targets with durations of zero, performance for which is, by definition, zero (cf. Equation 10). Overall, the effects of both stimulus duration and stimulus contrast were as expected: Performance increases with increases in each.

To address liftoff durations—the major issue we set out to investigate in Experiment 1—we examined performance values for the shortest study exposure durations. As indicated above, our theory generally predicts that with high contrast, liftoff is at or close to zero, which implies that performance will exceed zero at even very short durations—that is, 5, 10, and 20 msec. With low contrasts, however, the presence of a threshold engenders an above-zero liftoff duration; that is, performance for short study durations should be zero as Loftus and his colleagues found using low-contrast digit stimuli.

It is evident that this predicted data pattern emerged. For high-contrast stimuli, performances for the 5-, 10-, and 20-msec exposure duration conditions were substantially above zero (performance values of .250 ± .025, .523 ± .025, and .629 ± .025, respectively). On the other hand, corresponding performance levels for low-contrast stimuli were very close to zero (performance values of .015 ± .026, .002 ± .026, and .021 ± .026 respectively).

It is evident from inspection of Figure 8 that the "liftoff value" for the low-contrast stimuli is estimated to be approximately 20 msec, just as it was reported to be by Busey and Loftus (1994) and by Loftus and Ruthruff (1994) for digit recall when the digits were of approximately the same contrast level (approximately 3%). This is a remarkable commonality, as it suggests that the visual system treats any incoming stimulus—be it as simple as a digit string, or as complicated as a complex picture—the *same* with respect to contrast level.

Theory Fit

Thus, at a qualitative level, the liftoff effects are as predicted by the theory. We fit the quantitative version of

Table 1
Parameter Values and Root Mean Square Errors

				Parameter			
	Variables	n	τ	c	Θ	Y	RMSE
All experiments simultaneously		8	8.5	2	.04	.79	.064
Experiment 1	Stimulus duration × high/low contrast	8	5.0	2	.010	.78	.064
Experiment 2	Gap duration × high/low contrast	8	14.0	2	.02	.80	.034
Experiment 3	Gap duration × stimulus duration	8	6.0	2	.03	.80	.014
Experiment 4	Gap duration × masking	8	9.0	2	.02	.80	.031

Note—In Experiment 1, there were six stimulus durations, varying from 5 to 200 msec. In Experiments 2–4, there were four gap durations, varying from 0 to 200 msec. In Experiment 3, there were two total stimulus durations of 60 and 100 msec. In Experiment 4, the mask delay was 0 or 250 msec.

the theory described above to the Experiment 1 data. To reiterate, the theory incorporates five parameters: n, τ, Θ, c, and Y. We set n to 8, a value that had characterized a variety of past theoretical fits; hence there were four free parameters. We carried out grid searches to determine the parameter values that minimized the root mean square error (RMSE) between the theoretical and observed values.

In the estimation of parameter values, it was not clear a priori how we should treat the different subject groups in the different experiments. Two extreme possibilities are the following. First, one might suppose that the different subject groups were characterized by generally different parameter values, in which case it would be appropriate to estimate separate parameter values for the individual experiments. The other extreme possibility is that all subject groups are characterized by a single set of parameter values, in which case it would be appropriate to estimate this single set across all experiments. Naturally, other possibilities exist; for instance, some parameters may vary across subject groups, whereas others remain constant.

Lacking a clear resolution to this issue, we carried out the two kinds of grid searches corresponding to the two extreme possibilities. Using the *individual-fit procedure*, we applied the theory to each experiment individually (which, of course, resulted in somewhat different estimated parameter values for the individual experiments). The predicted values resulting from this procedure are represented by the dashed lines in Figure 8 (and in subsequent data figures). Second, using the *simultaneous-fit procedure*, we fit the theory simultaneously to all four experiments that we report in this article. As will be described in more detail below, there was a total of 36 experimental conditions across the four experiments. The predicted values resulting from this procedure are represented by the solid lines in Figure 8 (and in subsequent data figures). All resulting parameter values and RMSEs are provided in Table 1.

We will defer comments about the simultaneous-fit results until after we have described all four experiments. It is evident that the fits for Experiment 1 are reasonable. The major flaw is that the long-duration, low-contrast data are predicted to be too high. The main reason for this deficiency is that the two curves appear to be reaching dif-

ferent asymptotes, and we decided to include only a single asymptote in the quantitative version of the theory (thus in this and many other respects the quantitative version of the theory is incomplete; however, as we have already indicated, a complete quantitative theory is not our goal here). In any event, the reasonable quality of the fit that is obtained indicates that the theory—originally applied to simple digit recall data—can provide a reasonable quantitative account of contrast and duration variation effects in picture memory as well.

Experiment 2
Contrast/Gap Duration

The major goal in Experiment 2 was to investigate the gap effect under conditions of high and low contrast. We did so by measuring recognition performance for pictures shown for a total duration of 60 msec, divided into two temporal halves of 30 msec apiece, separated by a blank temporal gap that ranged from 0 to 200 msec. As indicated above, the most straightforward application of our theory to picture recognition implies that under low-contrast conditions, performance should decline with gap duration, as was observed in prior digit recall experiments. Under high-contrast conditions, the theory predicts a close-to-zero gap effect.

Method

Experiment 2 entailed the same general subject pool, materials, apparatus, and procedures as Experiment 1, with the following exceptions.

Subjects. Eighty subjects were run in sixteen 5-subject groups.

Materials. The materials were identical to those in Experiment 1 except that stimulus luminances and contrasts differed. In particular, luminance of the black and the white regions were measured on a sample slide. In the high-contrast condition, luminance (in cd/m^2) of the black regions was 58.7 and of the white regions 68.4, yielding a contrast of 7.6% [(w − b)/(w + b)]. In the intermediate (test phase) contrast condition, mean luminance was 57.1 and 61.5 for the black and the white regions, respectively, yielding a contrast of 3.7%. In the low-contrast condition, mean luminance was 56.4 and 58.5 for the black and the white regions, respectively, yielding a contrast of 1.8%. The adapting field alone was measured to be 54.9 cd/m^2.

Apparatus. The apparatus was the same as that used in Experiment 1. Again, the adapting field was present throughout; stimuli
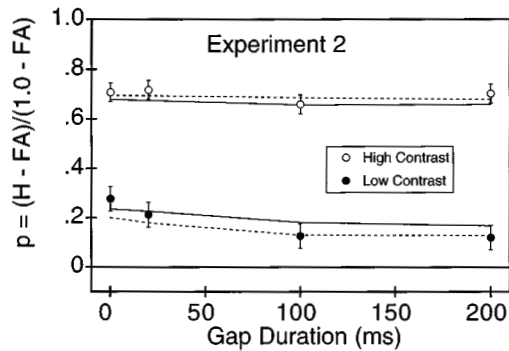
Figure 9. Experiment 2 data. Solid and dashed lines show theoretical fits.

presented at both study and test were superimposed over the adapting field.

Design. There were two independent variables in Experiment 2. The first was stimulus contrast, which was either high or low. The second was the duration of the gap that intervened between the first and second presentation of each stimulus; gap duration ranged from 0 to 200 msec. Both contrast and gap duration were varied within subjects. Gap duration was varied randomly across study trials within each study phase. As in Experiment 1, equipment limitations precluded this and required that contrast be varied across slide trays and remain constant within each slide tray.

Procedure. As in Experiment 1, an experimental session consisted of a study phase followed by a test phase for each of the four trays of slides. The specific procedures were as follows.

Study procedures. Within each stimulus tray, 24 stimuli were presented during the study phase at either low or high contrast. On each trial, the stimulus was presented for 30 msec, followed by a blank temporal gap (during which time only the adapting field was present) followed by a second 30-msec presentation; hence total exposure duration was 60 msec. Gap duration was 0, 20, 100, or 200 msec. The four gap duration conditions were presented in random order with the restriction that each gap duration occur twice within each of the three, eight-trial blocks. As in Experiment 1, contrast was blocked by tray.

Test procedures. Experiment 2 test procedures were identical to those of Experiment 1.

Counterbalancing. As in Experiment 1, target–distractor counterbalancing was accomplished as follows. Within each slide tray, the 48 slides were randomly divided into two sets (Set A and Set B) of 24 slides per set. For the eight odd numbered groups, Set A slides served as targets while Set B slides served as distractors; the reverse was true for the remaining eight (even numbered) groups. Each stimulus occurred once in each of the eight conditions defined by four gap durations $\times$ two contrast levels over the eight groups in which it appeared as a target.

As in Experiment 1, Trays 1–4 were presented at high, low, low, and high contrast levels, respectively (HLLH configuration), or low, high, high, and low, respectively (LHHL configuration). Half the even numbered groups and half the odd numbered groups received each of these two configurations.

Results and Discussion

The false alarm probabilities were .177 and .336 for the high- and low-contrast pictures. Performance for each of the eight gap duration $\times$ contrast conditions was computed using Equation 11. Figure 9 shows the main results of Experiment 2. It is evident that with low-contrast stimuli, picture recognition performance declines with gap duration, mimicking the digit recall data. With high-contrast stimuli, however, as predicted, there is virtually no gap effect. Table 2 quantifies these assertions, showing, for Experiments 2–4, a variety of gap effect assessments. The information in the top two rows of Table 2 provides several comparisons between the high-contrast and low-contrast comparisons for Experiment 2. First, the SSB column lists the sums of squares due to the gap manipulation: As can be seen, there is almost nine times as much performance variance engendered by gap variation in the low-contrast as opposed to the high-contrast condition. The column labeled SSB (Linear) provides sums of squares due to a linear contrast relating performance to gap duration (linear weights = 4, 3, -1, -6 for the gap durations of 0, 20, 100, and 200 msec). Here the difference between the high- and low-contrast conditions is even more pronounced: The linear variance component for the low-contrast condition exceeds that for the high-contrast condition by a factor of more than 60. Finally, the rightmost column of Table 2 indicates that the performance difference between the 0-msec and 200-msec gap conditions is substantial in the low-contrast condition (.158 $\pm$ .071) but minimal in the high-contrast condition (.007 $\pm$ .054).

In summary, the results of Experiment 2 confirm our quasi-qualitative predictions quite well. Figure 9 along with Table 2 indicates that both kinds of quantitative fits are also quite good.

Experiment 3
Stimulus Duration/Gap Duration

There is a potential alternative explanation for the lack of gap effect in the high-contrast Experiment 2 data which

Table 2
Comparisons of Gap Effects Under Different Experimental Circumstances

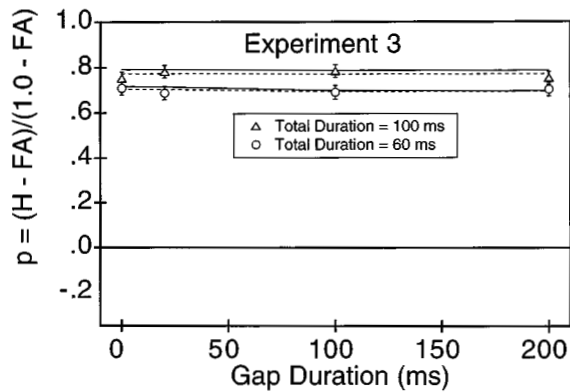|		Gap Effect Measure				
Experiment	Variable	SSB	SSB (Linear)	p_0 0-msec Gap	p_{200} 200-msec Gap	$p_0 - p_{200}$ ($\pm$95% CI)
2	High contrast	0.119	0.013	.708	.701	.007 $\pm$ .054
	Low contrast	1.021	0.791	.278	.120	.158 $\pm$ .071
3	100-msec total duration	0.056	0.003	.753	.753	.000 $\pm$ .040
	60-msec total duration	0.019	0.000	.710	.703	.007 $\pm$ .042
4	Delayed mask	0.013	0.008	.704	.695	.009 $\pm$ .054
	Immediate mask	6.568	6.029	.145	.564	$-$.419 $\pm$ .062

Figure 10: Experiment 3 data. Solid and dashed lines show theoretical fits.

involves a form of ceiling effect. It is most easily conceptualized as follows.

As we have discussed earlier, picture recognition performance may be broadly construed as being dependent on two sets of processes. Perceptual processes generate what may be thought of as raw perceptual information. Conceptual processes operating on this perceptual information provide a long-term representation of the stimulus picture which, in turn, provides a direct basis for the eventual recognition response.

It is likely that each of these processes has asymptotic limits. In particular, it is likely that no matter how much perceptual information is acquired from a picture in our experiments, a variety of factors—most notably (1) lack of adequate opportunity for conceptual processing and (2) target–distractor similarity—place an upper, less than perfect limit on eventual recognition performance. Accordingly, one could argue that the lack of gap effect for the high-contrast pictures of Experiment 2 occurs simply because all four gap conditions are at this limit.

To investigate this possibility, we replicated the four high-contrast Experiment 2 conditions but included an additional four conditions in which total exposure duration was increased from 60 to 100 msec, thereby allowing additional perceptual processing to occur. If the ceiling explanation is correct, such additional exposure duration could not lead to a performance increase.

Method

Subjects, Design, and Procedure. Ninety-six subjects were run in sixteen, 6-subject groups. The materials were the same as in Experiments 1–2. As in all experiments, an experimental session consisted of a study phase followed by a test phase for each of the four trays of slides. The specific procedures were as follows.

Study procedures. As in the previous experiments, 24 stimuli were presented during the study phase within each stimulus tray. On each trial, the stimulus was presented in two equal exposure durations (30–30 msec or 50–50 msec) with an intervening gap of duration 10, 20, 100, or 200 msec. Unlike in Experiments 1–2, contrast was not manipulated but instead was fixed at a high level. As indicated, total stimulus exposure duration was also manipulated: On each trial, a stimulus was presented for a total of 30 + 30 =

60 msec or 50 + 50 = 100 msec. The combination of four gap durations and two stimulus durations thus defined one of the eight experimental conditions into which a given stimulus fell. These eight conditions were presented during each 24-trial study phase in random order with the restriction that each condition occur once within each of the three, 8-trial blocks. All conditions were thus both within subjects and within slide trays.

Test procedures. These were identical to Experiment 1. Contrast was the same as in the study phase (i.e., high).

Counterbalancing. Each stimulus occurred once in each of the eight conditions defined by four gap durations × 2 durations over the eight groups in which it appeared as a target.

Results and Discussion

Because all eight conditions were randomly intermingled in the study phase of Experiment 3 (as compared with Experiments 1–2, wherein contrast varied across study–test blocks), there was only a single false alarm probability, which was .161. The main results of Experiment 3 are shown in Figure 10. Two findings are apparent. First, there is virtually no effect of gap duration for either of the two stimulus duration conditions (see Table 2, rows 3 and 4 for detailed support of these assertions). Second, there was an overall performance difference of .070 ± .021 between the two durations. This rules out the ceiling effect explanation for the high-contrast data of Experiment 2. As is indicated in Figure 10 and Table 1, the quantitative fit of the theory to the data is quite good.

Experiment 4
Gap Duration/Masking

In Experiments 1–3, the stimuli were not preceded or followed by masks of any sort. This was done for simplicity: It allowed us test the theory without having to burden it with masking assumptions.

However the theory, or earlier variants of it, have successfully accounted for experimental paradigms in which simple digit stimuli have been followed by a high-luminance, high-contrast noise mask (Loftus et al., 1993; Loftus et al., 1992). These accounts incorporated the simple assumption that a noise mask halted information acquisition at the instant of mask occurrence (cf. Sperling, 1964b).

In the course of running Experiments 1–3, we began to informally explore masking effects by presenting our stimuli followed by masks in various temporal configurations. We discovered, (somewhat to our surprise) that when high-contrast stimuli were followed by a mask in a gap duration paradigm, longer gap durations appeared to produce more visible stimuli.[5] An investigation of our theory's predictions about this phenomenon indicated why (according to the theory) this should happen.

The general idea is shown in Figure 11, which depicts sensory response functions in four circumstances. The left panels show zero-gap conditions and the right panels show 100-msec gap conditions. The dashed and solid vertical lines in the right panels depict stimulus offsets and second-presentation onset, respectively. The lower panels depict an immediate mask—that is, a mask presented im-
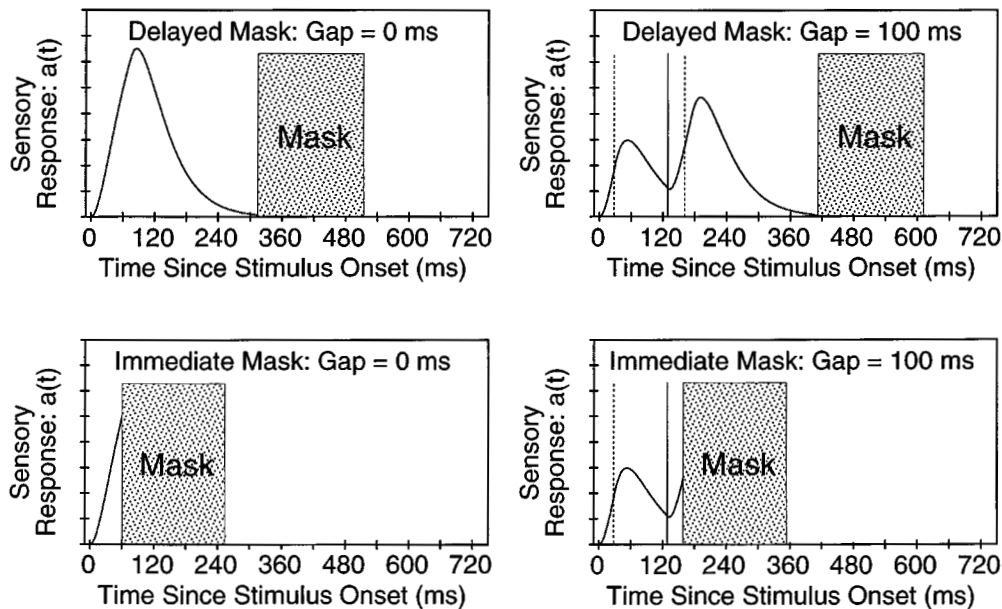
Figure 11: Effects of delayed masks (upper panels) and immediate masks (lower panels) on zero-gap presentations (left panels) and 100-msec presentations (right panels). The dashed and solid vertical lines in the right panels depict stimulus offsets and second-presentation onset, respectively.

mediately on offset of the second stimulus half. The upper panels show *delayed-mask* control presentations—that is, presentations that involve a mask (so as to control for any nonperceptual masking effects) but a mask that is delayed by 300 msec following stimulus offset so as to interfere only minimally with the sensory response.

As indicated, we assume that processing stops [i.e., that $r(t)$ falls to zero] at the instant of mask appearance. Thus, performance is determined by the area under the sensory response function prior to the mask. It is clear that, when the mask is delayed, it has virtually no effect: The unmasked areas under the two top-row sensory-response functions are the same (by the identity property). However, things are quite different with an immediate mask. When there is zero gap, the mask occurs immediately following the offset of a continuous 60-msec stimulus; hence no processing ever occurs when the stimulus is off. However, when there is a gap, processing can occur following offset of the first stimulus half, and the longer the gap, the longer such processing can progress. Accordingly, the longer the gap, the better the eventual performance.

Method

In Experiment 4, gap duration was varied as in Experiments 2 and 3. Only high-contrast stimuli were used. The second presentation on each trial was followed by a high-contrast noise mask that had been used in previous picture-memory experiments (e.g., Loftus, Johnson, & Shimamura, 1985). The mask occurred either immediately following stimulus offset (*immediate-mask* conditions), or 250 msec following stimulus offset (*delayed-mask* conditions). The delayed-mask conditions were designed to control for any masking effect on conceptual processing. Thus a mask always oc-

curred following a picture at study; what varied was whether it interfered or did not interfere with perceptual processing.

Subjects, Design, and Procedure. Forty subjects were run in eight, 5-subject groups. The materials were identical to those used in the previous experiments. Again, an experimental session consisted of a study phase followed by a test phase for each of the four trays of slides. The specific procedures were as follows:

Study procedures. Within each stimulus tray, 24 stimuli were presented during the study phase. On each trial, the stimulus was presented for 30 msec, followed by a gap followed by a second 30-msec presentation. Gap duration was 10, 20, 100, or 200 msec. Contrast was fixed at a high level. A noise mask presented for 200 msec followed the second 30-msec stimulus presentation either immediately or after a 250-msec blank delay. The combination of four gap durations and immediate/delayed mask thus defined one of the eight experimental conditions into which a given stimulus fell. These eight conditions were presented in random order with the restriction that each condition occur once within each of the three eight-trial blocks. All conditions were both within subjects and within slide tray.

Test procedures. These were identical to those for Experiments 1 and 3. Test picture contrast was high.

Counterbalancing. Each stimulus occurred once in each of the eight conditions defined by four gap durations × two masking conditions over the eight groups in which it appeared as a target.

Results and Discussion

As in Experiment 3, all eight conditions were randomly intermingled in the study phase of Experiment 3 and there was thus only a single false alarm probability, which was .253. Figure 12 shows the main results of Experiment 4; see also Table 2 (rows 5–6) for a comparison of gap effects in the delayed- and immediate-mask conditions.

There was essentially no gap effect when the mask was delayed. In addition, *mean* delayed-mask performance
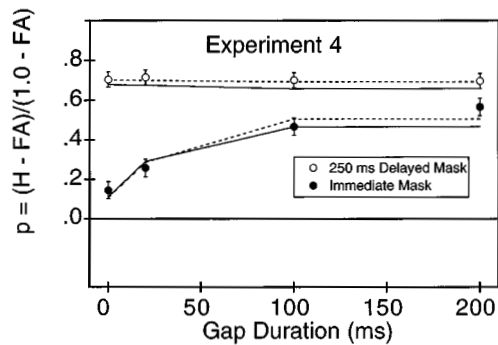
Figure 12. Experiment 4 data. Solid and dashed lines show theoretical fits.

(.704) was almost identical to corresponding high-contrast, 60-msec total duration performance in Experiments 2 and 3 (mean performance = .696 and .698, respectively). In short, a delayed mask had very little effect on eventual recognition performance.

When the mask was immediate, however, there was a strong gap effect: It is evident from Figure 12 that, consistent with our informal observations and theoretical analyses, immediate-mask performance increased dramatically as gap duration increased from 0 to 200 msec.

All eight conditions were well predicted by the theory.

Interexperiment Issues

Before proceeding to our General Discussion, we briefly discuss two issues pertaining to the four experiments considered in concert: Inter-experiment consistency, and the overall theoretical fit.

Consistency

The four experiments that we have reported involved separate groups of subjects and were carried out at different times. Because we attempted to fit the same theory to all of them, the issue of interexperimental consistency is an important one. To the degree that confoundings involving subject populations and time of running are associated with different strategies, different theoretical parameters, or, more generally, different performance levels, it becomes less appropriate to fit all data at once.

The issue of consistency is best addressed by examining performance in condition sets that are relatively similar over experiments. Three such condition sets are (1) Experiment 2, high contrast; (2) Experiment 3, 60-msec total duration; and (3) Experiment 4, delayed mask. Note that these condition sets are not *identical*; they differed in minor respects. In Experiment 2, only gap duration varied within a study phase, whereas in Experiment 3, both gap duration and stimulus duration varied within a study phase. In Experiment 4, a delayed noise mask followed stimulus offset, and both delayed- and immediate-mask condition sets occurred within a study phase.

Table 3 shows performance for these three condition sets. We have already noted that mean performance is very similar (see the rightmost column); we note that performance is likewise similar over the three condition sets for each of the four individual gap duration levels.

Thus in all instances wherein we are able to compare across experiments, the corresponding data are very similar. This observation provides us with some degree of confidence that confoundings entailed in using different subjects and running the experiments at different times did not materially affect our results.

Overall Theoretical Fit

As we have indicated, the solid lines in Figures 8–10 and 12 represent the best fit of four free experimental parameters (i.e., those listed in Table 1 minus the linear-filter parameter, n, which was set at eight a priori) to the 34 total experimental conditions across the experiments. The overall RMSE was .064 or approximately 7% of the total range of the data points.

The bulk of the theory–data discrepancy lies in Experiment 1, where the low-contrast data are somewhat worse than predicted. Given the high degree of interexperiment consistency that we found, this means that either (1) the theory cannot adequately account simultaneously for stimulus duration in conjunction with the other variables—stimulus contrast, gap duration, and immediate/delayed mask—across our experiments, or (2) some factor beyond the theory's domain is causing the Experiment 1 discrepancies. As we have already discussed, it appears that one obvious theoretical inadequacy is our assumption of only a single asymptote, approached by both low- and high-contrast conditions. However, because we do not aspire here to a perfect quantitative fit, we have not pursued this issue.

GENERAL DISCUSSION

Summary

Our major goal in this article has been to show how a theory that has been successful in describing the relation between sensory processing and information acquisition in a fairly simple situation—immediate digit recall—can be used to play an analogous role in the more complex situation wherein complex, naturalistic pictures are tested in a later recognition procedure. A tentative conclusion is that intertwined sensory and information-acquisition processes are undertaken by the visual system without

Table 3
Comparable Conditions in Three Experiments

Experiment	Variable	Gap Duration (msec)				
		0	20	100	200	M
2	High contrast	.708	.717	.659	.701	.696
3	60-msec total duration	.710	.688	.691	.703	.698
4	Delayed mask	.704	.715	.700	.695	.704

regard to stimulus complexity or to the eventual task for which the acquired information is to be used—and that our theory captures these processes.

In Experiments 1–3, with low picture contrast, we replicated two nonobvious effects—the liftoff effect and the gap effect—predicted by our theory, which have been observed in low-contrast, digit recall tasks. With higher picture contrast, two other nonobvious, but predicted consequences were confirmed: Both the liftoff and the gap effect largely disappear. Finally, in Experiment 4, we applied our theory to a situation in which a high-contrast noise mask followed the stimulus. Our assumptions about masking's effect was exceedingly simple—that the mask simply served to terminate information acquisition. Nonetheless, with no other modification (and no alteration in existing parameter estimates), the theory accounted for the rather complex pattern of effects extremely well.

Theoretical Simplification and Quantitative Predictions

The theory that we have described represents the conjunction of two basic theoretical components, both of which are simplified and incomplete. The linear-filter front end has been used by vision scientists to describe a variety of psychophysical data under conditions of central viewing, using very weak (near threshold) stimuli. The information-acquisition assumptions have likewise been applied in the past to stimuli that are simple as cognitive stimuli go—generally alphanumeric characters—and these assumptions pay little if any attention to higher level characteristics of the stimuli, such as their organization or semantic content. To these limitations of the fundamental theoretical components, we have added another simplification in order to apply the theory to picture memory: We have essentially eliminated a recognition decision mechanism and replaced it with the weak assumption of a monotonic function relating acquired perceptual information and eventual recognition performance.

Even with these vast simplifications, the theory has been remarkably successful in accounting both qualitatively and quantitatively for the effects on picture-recognition performance of a number of independent variables, including exposure duration, stimulus contrast, gap duration, and whether the stimulus was followed by a noise mask or by a blank field. Of particular note is that the observed value of "liftoff" (the minimum duration to produce above-chance performance) is the same (approximately 20 msec) for a specific contrast level (approximately 3% contrast) whether this value is measured by immediate digit recall (Busey & Loftus, 1994; Loftus & Ruthruff, 1994) or by delayed picture recognition as in the present Experiment 1. This commonality presumably reflects a fundamental property of the visual system in terms of how it treats weak, briefly presented stimuli.

Despite our theoretical condensation of important sets of processes, we undertook the task of devising one quantitative implementation of our theory. This undertaking was meant to serve as a demonstration that there indeed exists at least one such quantitative implementation that will provide a reasonable quantitative fit to our data. In this quantitative model, we assumed, for example, the magnitude of our memory measure in a given condition to be equal to the amount of acquired information in that condition. This is obviously an extremely simplistic assumption, but it works reasonably well nonetheless, and it therefore follows that an accurate, but more complete account of the recognition process (e.g., Murdock, 1993) would make predictions that would be approximately the same, but would be derivable from more fundamental decision processes.

Information Extraction and Phenomenological Appearance

In this article, we have focused on one ecologically important facet of visual behavior: acquisition of information from a stimulus. Another similarly important facet of visual behavior is the *phenomenological appearance* of the stimulus. We have thus far alluded only briefly to phenomenological appearance, but our data and our theory provide a context for considering it and, accordingly, we will briefly discuss it in this final section.

In the high-contrast conditions of Experiments 2–4, there was generally no effect of gap duration on performance (see Tables 2 and 3). However, as we suggested earlier, informal observations (which were subsequently verified experimentally and will be reported in a forthcoming article) indicate that it is trivially easy at the time of study to distinguish, for example, whether a stimulus is in the 0-msec or the 200-msec gap condition. In other words, information about gap duration that is clearly available to an observer at the time of original perception does not affect subsequent recognition-memory performance.

What are we to make of this apparent dissociation between information acquisition (as reflected by later memory performance) on the one hand and phenomenological appearance on the other? As described in more detail by Loftus and Hogden (1988), Loftus and Ruthruff (1994), and Loftus and Irwin (1998), information acquisition and phenomenological appearance can profitably be conceptualized as reflecting two properties of the same internal function—the information acquisition rate function, $r(t)$. In particular we have, in past articles, made the assumption that the degree to which a stimulus appears to be phenomenologically present at any given time is determined by the degree to which the observer is processing the stimulus—that is, by the magnitude of $r(t)$—at that time. This assumption has permitted us to account for the inverse duration effect in synchrony judgment experiments (e.g., Efron, 1970a, 1970b) and in temporal integration tasks (e.g., Di Lollo, 1980; Hogben & Di Lollo, 1974). Given this assumption, the *shape* of the $r(t)$ function represents the temporal course of phenomenological appearance, while the *area under* the $r(t)$ function represents, by

definition, the amount of information acquired from the stimulus. This hypothesis and its implications are, at present, major foci of investigation in our laboratory.

REFERENCES

Biederman, I. (1993). Geon theory as an account of shape recognition in mind and brain. *Irish Journal of Psychology*, **14**, 314-327.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review*, **101**, 446-469.

Busey, T. A., & Loftus, G. R. (1998). Binocular information acquisition and visual memory. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1188-1214.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, **109**, 75-97.

Di Lollo, V., & Bischof, W. F. (1995). Inverse-intensity effect in duration of visible persistence. *Psychological Bulletin*, **118**, 223-237.

Di Lollo, V., & Dixon, P. (1992). Inverse duration effects in partial report. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 1089-1100.

Edelman, S., & Poggio, T. (1990, April). *Bringing the grandmother back into the picture: A memory-based view of object recognition* (MIT Artificial Intelligence Memo No. 1181). Cambridge, MA: MIT.

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

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

Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.

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

Hinton, G. E. (1981). A parallel computation that assigns canonical object-based frames of reference. In A. Drinan (Ed.), *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*. Vancouver, BC: IJCAI.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528-551.

Hintzman, D. L., Curran, T., & Caulton, D. A. (1995). Scaling the episodic familiarity of pictures and words. *Psychological Science*, **5**, 308-313.

Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research*, **14**, 1059-1069.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, **99**, 480-517.

Hummel, J. E., Biederman, I., Gerhardstein, P. C., & Hilton, H. J. (1988). From image edges to geons: A connectionist approach. In D. Touretsky, G. Hinton, & T. Sejnowski (Eds.), *Proceedings of the 1988 Connectionist Models Summer School* (pp. 462-471). San Mateo, CA: Morgan Kaufman.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 1-12.

Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 115-125.

Lindsay, P. H., & Norman, D. A. (1977). *Human information processing: An introduction to psychology* (2nd ed.). New York: Academic Press.

Loftus, G. R., & Busey, T. A. (1992). Multidimensional models and iconic decay. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 356-361.

Loftus, G. R., Busey, T. A., & Senders, J. W. (1993). Providing a sensory basis for models of visual information acquisition. *Perception & 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 & Performance*, **18**, 530-549.

Loftus, G., R., Hanna, A. M., & Lester, L. (1988). Conceptual masking: How one picture captures attention from another picture. *Cognitive Psychology*, **20**, 237-282.

Loftus, G. R., & Hogden, J. (1988). Picture perception: Information extraction and phenomenological appearance. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 139-191). New York: Academic Press.

Loftus, G. R., & Irwin, D. E. (1998). On the relations among different measures of visible and informational persistence. *Cognitive Psychology*, **35**, 135-199.

Loftus, G. R., Johnson, C. A., & Shimamura, A. P. (1985). How much is an icon worth? *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 1-13.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, **1**, 476-490.

Loftus, G. R., & Ruthruff, E. R. (1994). A theory of visual information acquisition and visual memory with special application to intensity-duration tradeoffs. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 33-50.

Lowe, D. G. (1987). The viewpoint consistency constraint. *International Journal of Computer Vision*, **1**, 57-72.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Metcalfe-Eich, J. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89**, 609-626.

Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review*, **100**, 183-203.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 509-522.

Raaijmakers, J. G., & Shiffrin, R. M. (1992). Models for recall and recognition. *Annual Review of Psychology*, **43**, 205-234.

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

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In *Symposium on the mechanism of thought processes*. London: Her Majesty's Stationary Office.

Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American*, **203**, 60-68.

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 591-600.

Snodgrass, J. G., & Feenan, K. (1990). Priming effects in picture fragment completion: Support for the perceptual closure hypothesis. *Journal of Experimental Psychology: General*, **119**, 276-296.

Sperling, G. (1964a). Linear theory and the psychophysics of flicker. *Documenta Ophthalmologica*, **18**, 3-15.

Sperling, G. (1964b). What visual masking can tell us about temporal factors in perception. In *Proceedings of the Seventeenth International Congress of Psychology* (pp. 199-200). Amsterdam: North-Holland.

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

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.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, **32**, 193-254.

Wasserman, G. S. (1991). Time and duration: A persistent illusion. *Perception & Psychophysics*, **50**, 603-604.

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 to 6-43). New York: Wiley.

NOTES

1. Both object recognition and recognition decision processes have been studied in detail, and sophisticated models of these processes have been developed. Existing object recognition models include template-matching models (e.g., Edelman & Poggio, 1990; Lowe, 1987; Ullman, 1989), feature list matching models (Hinton, 1981; Hummel, Biederman, Gerhardstein, & Hilton, 1988; Lindsay & Norman, 1977; Selfridge, 1959; Selfridge & Neisser, 1960), and structural representation models (e.g., Hummel and Biederman, 1992). Several recognition memory models also exist that address processes that occur at the time of retrieval and effects of the target–distractor relationship (e.g., Gillund & Shiffrin, 1984; Hintzman, 1986, 1988; Metcalfe-Eich, 1982; Murdock, 1982, 1993). In our theory, we do not commit to particular models of object recognition or recognition decision. We simply subsume these processes in a simple fashion within a theory that describes in detail the first stages of picture perception and memory.

2. Although it will not be critical for the present experiments, a parenthetical note is in order. The representation in Figure 1A presumes that the stimulus is presented by some device such as a tachistoscope that displays the stimulus at a constant luminance value throughout the stimulus's physical presence. Such would not be true if the stimulus were presented—as is often the case—by a raster-scan CRT. Busey and Loftus (1994) analyzed such a situation and concluded that the theory's predictions for a constant-display device and a raster-scan device are equal to well within experimental noise.

Several comments about our notational scheme are in order. First, t is time since stimulus onset, while d refers to the experimentally set stimulus duration. We emphasize this distinction, as t and d are often confused (cf. Wasserman, 1991). Second, when the stimulus temporal waveform is a step function (as in Figure 1A) the constant contrast level is denoted Φ (thus in Figure 1, $\Phi = 0.05$).

3. Such specific quantitative theoretical predictions are not, of course, possible if the relation between acquired information, $I(\infty)$, and performance, p, were assumed to be merely monotonic. To obtain these predictions, we strengthened the assumed monotonic function to the identity function, $p = I(\infty)$, for reasons that are provided in the relevant references. It is important to note, however, that the liftoff effect (defined as a liftoff greater than zero) is predicted by the theory no matter what the nature of the assumed monotonic function. The same is true with the contrast effects and the "gap effect" to be described in the next two sections.

4. Our assumptions about picture memory appear to carry with them the dubious implication that information acquired from a picture is *unidimensional*—that it can be represented by a single number on a unidimensional scale. For example, reference to a monotonic relation between "degree of conceptual processing and recognition memory performance" (or reference to "degree of conceptual processing" to begin with, for that matter) implies a unidimensional scale that corresponds to "degree of conceptual processing." We note that this assumption has some degree of empirical support: For instance, Hintzman, Curran, and Caulton (1995), using multidimensional scaling techniques, showed that a single unidimensional scale was sufficient to account for both recognition memory and frequency judgments in picture and word memory.

In the present article, we do not explicitly define what these scales are, although it can be done in a variety of ways (e.g., Loftus et al., 1988). Nonetheless, few would seriously consider such assumptions to be viable; it is almost self-evident that pictures, along with the sorts of information about them that humans acquire, are complex and multidimensional.

This issue has been considered in the past with respect to the present theory (Di Lollo & Dixon, 1992; Loftus & Busey, 1992; Loftus et al., 1992). Briefly, its resolution is this. The information acquired from a picture may be represented as a point in J-dimensional space (where J is the number of informational dimensions in a picture). The process of acquiring information may be represented as tracing out a path in this space (presumably beginning at the origin and always moving in a nonnegative direction along all J dimensions). The distance along this path from the origin to the current point at any given time, t, *is* unidimensional, and is what is reflected by $I(t)$.

5. Low-contrast stimuli appeared to be essentially invisible when followed by a mask.