Journal of Experimental Psychology:
Human Perception and Performance
2004, Vol. 30, No. 1, 104–118

0096-1523/04/$12.00 DOI: 10.1037/0096-1523.30.1.104

How Different Spatial-Frequency Components Contribute to Visual Information Acquisition

Geoffrey R. Loftus and Erin M. Harley
University of Washington

We test 3 theories of global and local scene information acquisition, defining global and local in terms of spatial frequencies. By *independence theories*, high- and low-spatial-frequency information are acquired over the same time course and combine additively. By *global-precedence theories*, global information acquisition precedes local information acquisition, but they combine additively. By *interactive theories*, global information also affects local-information acquisition rate. We report 2 digit-recall experiments. In the 1st, we confirmed independence theories. In the 2nd, we disconfirmed both independence theories and interactive theories, leaving global-precedence theories as the remaining alternative. We show that a specific global-precedence theory quantitatively accounted for Experiments 1–2 data as well as for past data. We discuss how their spatial-frequency definition of spatial scale comports with definitions used by others, and we consider the suggestion by P. G. Schyns and colleagues (e.g., D. J. Morrison & Schyns, 2001) that the visual system may act flexibly rather than rigidly in its use of spatial scales.

Since the time of Neisser's (1967) classic *Cognitive Psychology*, visual scenes have been conceptualized as being decomposable into *global* and *local* information: Global information corresponds to overall scene structure, whereas local information corresponds to fine details (see Morrison & Schyns, 2001, pp. 454–456, for a summary). This article is about the relations between global information and local information—the relative time courses over which they are acquired and the means by which they combine into an overall perception of the scene.

Spatial-Frequency Information

Global and local information can be operationalized in many ways. One of them is in terms of *spatial frequencies.* From the visual system's perspective, the world is made up of different spatial frequencies. An example is provided in Figure 1, the top of which shows a typical real-world scene. The two bottom panels show the same picture filtered to pass only the low spatial frequencies (LSFs; bottom left) or high spatial frequencies (HSFs; bottom right). It is evident that the two spatial-frequency components carry different kinds of information that accord well with the general conceptualization of global information and local information: The bottom left Figure-1 picture carries a global representation of the scene, whereas the bottom right picture conveys information about edges and details, such as the writing on the dollar bill. There is indeed ample evidence that the visual system decomposes the visual scene into separate spatial-frequency components (Blakemore & Campbell, 1969; Campbell & Robson, 1968; De Valois & De Valois, 1980, 1988; Graham, 1989; Olzak & Thomas, 1986). Accordingly, a correspondence between global–local on the one hand and LSF–HSF on the other has been hypothesized by numerous investigators (e.g., Hughes, Nozawa, & Ketterle, 1996; Schyns & Oliva, 1994), and we shall use it in the remainder of this article.

Theories of the Relations Between Global and Local Information

We begin by defining and characterizing three classes of theories of the relations between LSF and HSF information. We first divide theories into two mutually exclusive classes: *independence* theories and *global-to-local* (GtL) theories. We then further subdivide GtL theories into *global-precedence* GtL theories and stronger, *interactive* GtL theories. For the moment, we characterize the distinctions informally and verbally. In later sections, we provide formal, quantitative instantiations of them.

Independence Theories

According to independence theories, both global information and local information are acquired in a manner that does not depend on when the information is acquired after stimulus onset (e.g., Soloman & Pelli's, 1994, single-filter model; see also Majaj, Pelli, Jurshan, & Palomares, 2002). To illustrate, suppose that one could manipulate when, after stimulus onset, the two types of information are acquired by presenting global information before local information or vice versa. By an independence theory, the order that the two types of information were presented would not matter (as found by, e.g., Parker, Lishman, & Hughes, 1996): The resulting total information would be the same. Independence the-

Geoffrey R. Loftus and Erin M. Harley, Department of Psychology, University of Washington.

Erin M. Harley is now at the Department of Psychology, University of California, Los Angeles.

This research was supported by National Institute of Mental Health Grant MH41637 to Geoffrey R. Loftus. We thank Liddy Olds, Tom Sanocki, and Philippe Schyns for helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Geoffrey R. Loftus, Department of Psychology, Box 351525, University of Washington, Seattle, WA 98195-1525. E-mail: gloftus@u.washington.edu

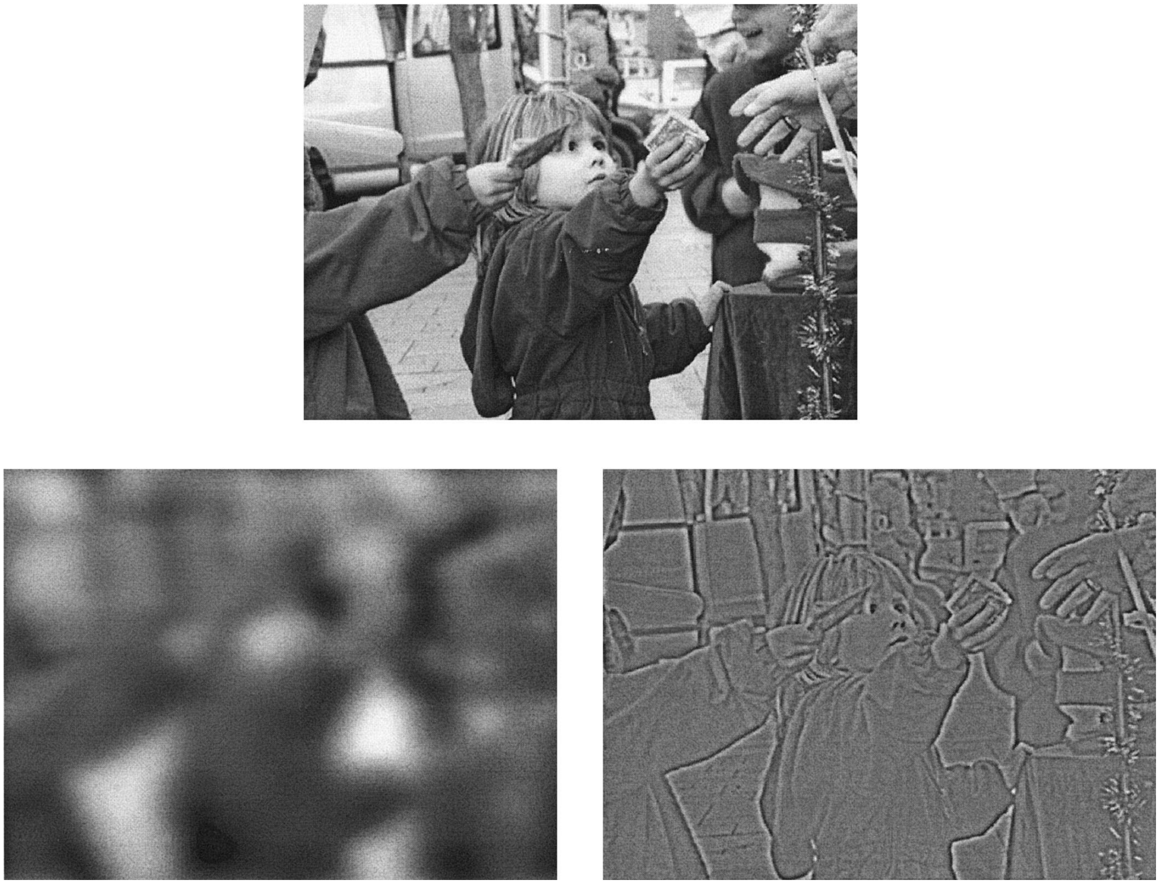

Figure 1. Decomposition of a naturalistic scene (top) into low-spatial-frequency components (bottom left) and high-spatial-frequency components (bottom right).

ories, moreover, assume *additivity*: LSF and HSF information are summed to arrive at the total information. An analogy is this. Suppose that you plan to buy vegetables from the vegetable stand and to buy candy from the candy store. It does not matter in which order you do these errands; when you arrive home you will have the same food in the same state.

Global-to-Local Theories

According to GtL theories, global information is acquired initially, followed by local information. Sanocki (2001) explicitly distinguishes between these two types of GtL theories. He has characterized acquisition-rate-difference theories by noting that, within their context, the time course of information acquisition differs for large-scale features (global information) versus small-scale features (local information): The former is acquired relatively quickly at first and then more slowly, but vice versa for the latter. In contrast, Sanocki (2001) has characterized interactive theories, which his data favor, by noting that within their context, "early large-scale information provides a large framework that can subsequently integrate small-scale information across the entire object" (p. 296). The distinction that we make here is much the same as that of Sanocki. For expositional convenience, we endow these two kinds of GtL theories with formal names: *global-precedence* GtL theories and *interactive* GtL theories.

In global-precedence GtL theories (e.g., Loftus, Nelson, & Kallman, 1983; Navon, 1977; Parker & Costen, 1999; Schyns & Oliva, 1994; Watt, 1987), acquisition of global information *precedes* acquisition of local information. Thus, the order in which the two types of information were presented would matter: Global information provided before local information would be better than the other way around. Like independence theories, global-precedence theories assume additivity of the LSF and HSF signals. To continue with our whimsical shopping analogy, suppose you go out planning to buy frozen vegetables at the supermarket and ice cream at the ice cream parlor, both of which you plan to eat on your arrival back home. Here, the order in which you buy your groceries matters: It is better to buy the frozen vegetables first and the ice cream last, because when you arrive home you want your vegetables to be thawed and your ice cream to be still frozen.

In interactive GtL theories (e.g., Navon, 1977; Sanocki, 1991, 1993, 2001), not only does global information precede local information but acquisition of local information also *depends on* the amount of already acquired global information. That is, acquired global information provides a spatial framework, within which local information can be interpreted and integrated—and the more complete the global information, the more efficient is such local processing. Thus, by interactive theories, LSF and HSF signals do *not* combine additively. Note that the interactive version of a GtL theory is a special case of the global-precedence version: The interactive version can be correct only if the global-precedence version is correct, but not vice versa.

Present Work

The research reported here was motivated largely by the results of two recent elegant experiments bearing on the time course over which HSF information and LSF information are acquired by the visual system and on the rules by which the two kinds of information combine to create an overall, coherent perception of a visual scene. We describe these experiments in some detail.

Schyns and Oliva (1994) created pictures that were composites of two entirely different scenes: one including only the LSF components and the other including only the HSF components. Observers viewed the composites either at a short duration (30 ms) or at a long duration (150 ms). At the short duration, observers had a strong tendency to perceive the LSF scene, whereas at the longer duration, they had a strong tendency to perceive the HSF scene. Schyns and Oliva concluded that LSF and HSF information operate over different time courses: LSF information is acquired first, followed by HSF information. Their data, by this conclusion, disconfirm independence theories and are consistent with either version of GtL theories. We note that, despite this finding, Schyns and his colleagues (Morrison & Schyns, 2001) have more recently retreated in several ways from the strong claim of a universal global–local processing order. We return to this topic and describe the rationale for this repositioning in our General Discussion.

Olds and Engel (1998) investigated acquisition of HSF and LSF information in several experiments involving identification of pictures of common objects. The pictures were of three types. *Intact* pictures were unaltered (except for being of quite low contrast). *LSF* pictures were created by low-pass filtering the intact pictures, thereby producing blurred versions akin to that shown in Figure 1, lower left. *HSF* pictures, like the one in Figure 1, lower right, were created by subtracting, on a pixel-luminance basis, the LSF picture from the intact picture and then adding the mean luminance of the intact picture. The intact picture was therefore the pixel-luminance mean of the LSF and HSF pictures. All three picture types were shown at exposure durations ranging from 17 to 100 ms, followed by a noise mask, and observers attempted to identify the objects depicted by them.

Olds and Engel (1998) tested whether performance (percentage of correct identification) for intact pictures could be accounted for by an independence theory. To do so, they fit to their data a theory called the *sensory response–information acquisition* (SRIA) theory, which is an instance of what we have defined to be an independence theory: The order of spatial-frequency components does not matter, and the spatial-frequency components combine additively. The SRIA theory has been described in detail elsewhere as it applies to perception and memory of digit strings (Busey & Loftus, 1994, 1998; Loftus, Busey, & Senders, 1993; Loftus & Ruthruff, 1994); line drawings (Loftus & McLean, 1999); and random forms (Harley, Dillon, & Loftus, 2003). To understand the meaning of Olds and Engel's results, it is necessary to understand this theory, so we describe it here briefly.

The theory begins with the physical stimulus, which is represented by a function, $f(t)$, relating stimulus contrast, C, to time, t, since stimulus onset. For stimuli presented on a CRT, as in Olds and Engel's (1998) (and the present) experiments, $f(t)$ consists of a series of pulses, each pulse corresponding to a single screen refresh. However, as shown by Busey and Loftus (1994, Appendix D), given the temporal parameters of the human visual system and the timing of CRT refresh rates, almost perfect predictions can be generated with the simplifying assumption that $f(t)$ is rectangular with a width equal to d, the stimulus duration, and a height equal to C, the stimulus contrast.

The stimulus is assumed to trigger a *sensory response* function, $a(t)$, which is the convolution of $f(t)$ and the system's *impulse-response* function, $g(t)$; thus,

$$a(t) = f(t) * g(t), \tag{1}$$

where * signifies convolution. The impulse-response function is a gamma function that is a convolution of n exponential decay functions, each with decay parameter τ; thus, n and τ are free parameters. At any post-sensory response stage, the system responds only to the magnitude by which $a(t)$ exceeds some *sensory threshold*, θ, a free parameter. Above-threshold sensory-response magnitude at time t is termed $a_\theta(t)$; more specifically,

$$a_\theta(t) = \begin{matrix} a(t) - \theta & \text{for } a(t) > \theta \\ 0 & \text{for } a(t) \leq \theta \end{matrix} . \tag{2}$$

At time t since stimulus onset, some proportion of information, $I(t)$, is assumed to have been acquired from the stimulus. New information is acquired at a rate proportional to the product of $a_\theta(t)$ and $[1 - I(t)]$. The proportionality constant is $1/c$, where c is a free parameter. Total acquired information from the stimulus, $I(\infty)$, can be shown to equal $1 - e^{A_\theta(\infty)/c}$, where $A_\theta(\infty)$ is the total area under $a_\theta(t)$. Performance, p, generally measured as some form of proportion correct (corrected for any guessing rate), is assumed equal to $I(\infty)$, the total information acquired from the stimulus over the course of stimulus presentation. The theory thus begins with the observable stimulus, $f(t)$ and, via parameters n, τ, θ, and c, generates observable performance, p.

To apply the SRIA theory to their experiment, Olds and Engel (1998) assumed that sensory response functions corresponding to the LSF and HSF components were *weighted*: The weight accorded the LSF component, w_L, was a free parameter, whereas the weight accorded the HSF component was $(1 - w_L)$.[1] They then assumed the LSF and HSF components simply added to produce the intact sensory response function. The SRIA theory fit Olds and Engel's data essentially perfectly, thereby confirming that an independence theory is sufficient to account for data issuing from spatial-frequency mixtures in this task.

Olds and Engel (1998) appeared to view this result as inharmonious with Schyns and Oliva's (1994) data, citing the latter as an example of data that "support the idea that responses to components at different scales combine non-linearly over the time-course

[1] Olds and Engel (1998) allowed two separate parameters, w_H and w_L, corresponding to HSF and LSF information. As we show in the Appendix, however, two weight parameters overdetermine the theory, and Olds and Engel's treatment is equivalent to that provided here with weights w_L and $w_H = (1 - w_L)$.

of object recognition” (p. 2109) and going on to characterize this result, along with those of Sanocki (1993), as implying “that the value of high spatial-frequency component depends on the amount of low spatial-frequency information that has already been processed” (p. 2010). Below, we take issue with this conclusion and argue that Schyns and Oliva’s data, although inconsistent with independence theories, do not imply an interactive theory. We also demonstrate that although Olds and Engel’s results are consistent with independence theories, they do not provide a sufficiently exacting test of independence theories; as we see, their data are also consistent with a global-precedence theory.

We have three goals in this article. The first is to replicate and extend Olds and Engel’s (1998) findings. The second is to modify Olds and Engel’s theory to account for new data that are inconsistent with it. The third is to demonstrate that a global-precedence theory accounts for all data under consideration.

Experiments

We report two experiments. In the first, we replicate and extend Olds and Engel’s (1998) Experiment-1 findings. In the second, we provide a more stringent test of an independence theory by investigating temporal rather than spatial relations between HSF and LSF components.

Experiment 1: Replication of Olds and Engel (1998)

Experiment 1 was designed to replicate and extend Olds and Engel’s (1998) Experiment-1 findings. Our experiment differs from theirs in three ways. The first two were implemented in quest of generalization: We used alphanumeric stimuli (digit strings) rather than pictures of objects as stimuli, and we used two different degrees of low-pass spatial filtering. The third difference was that we used unmasked in addition to masked stimuli. The reason for this, in addition to simply generalizing, was that interpretation of data using masked stimuli is often clouded because an auxiliary theory of masking must be appended to whatever main theory is being tested.

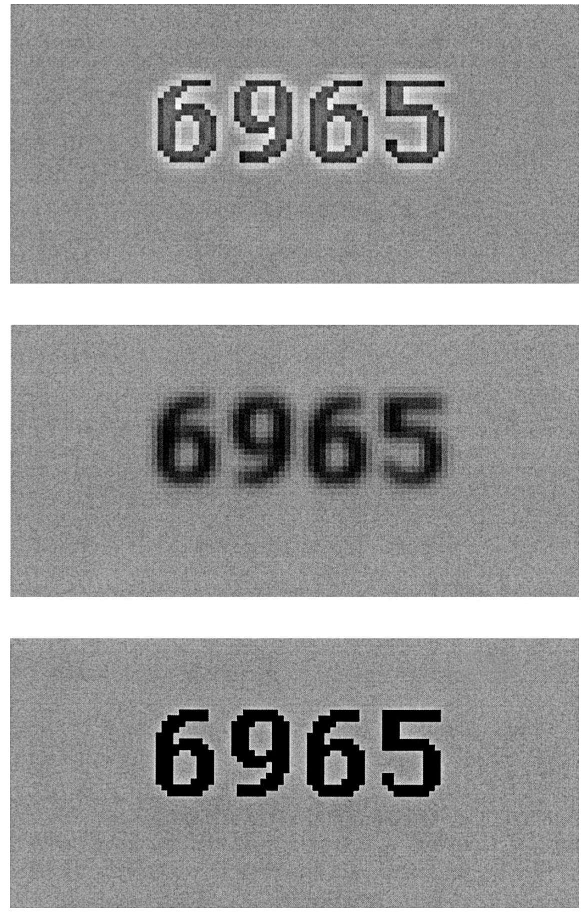

Figure 2. Examples of HSF (top), LSF (middle), and normal (bottom) stimuli. For ease of viewing, contrasts of all stimuli are considerably higher than they were in the experiments.

Method

Observers. Observers were 3 paid University of Washington undergraduates with normal or corrected-to-normal vision.

Apparatus. Stimuli were displayed on a 17-in. Macintosh ColorSync monitor with a refresh rate of 13.5 ms, driven by a Macintosh G3 computer. The experiment was executed in MATLAB using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The laboratory was dimly illuminated during the experiment. All stimuli were shown against a uniform gray background whose luminance was 8.3 cd/m^2. Observers’ eyes were approximately 18 in. (45.72 cm) from the screen.

Stimuli. Stimuli were randomly-generated four-digit strings, in Geneva font, displayed at a 20-pixel font height and subtending a vertical visual angle of approximately 0.9°.

As exemplified in Figure 2, stimuli appeared in three versions: intact, LSF, and HSF. An LSF stimulus was generated by Fourier-transforming an intact stimulus from pixel space to frequency space, multiplying the resulting amplitude spectrum by a half-Gaussian filter with a standard deviation of x cycles/digit (c/d) and inverse-Fourier-transforming the result back to pixel space. There were two values of x: 1.38 c/d (strong low-pass filtering, which cut off relatively more HSFs) and 1.95 c/d (mild low-pass filtering, which cut off relatively fewer HSFs). The HSF stimulus was generated by subtracting the LSF stimulus from the intact stimulus and then adding the intact stimulus’s mean luminance. Therefore for each degree of low-pass filtering, the luminance of the intact stimulus was the mean of the LSF stimulus and its HSF counterpart.

Stimulus contrast was defined as $(S - B)/(S + B)$, where S is stimulus luminance and $B = 8.3$ cd/m^2, is background luminance. Contrast for the intact stimuli varied across observers to compensate for their varying abilities and is shown in Table 1, Row 1, for each observer.

Design. Each observer participated in 96 data collection sessions, divided into four subexperiments of 24 sessions/subexperiment. The four subexperiments consisted of the four combinations of mild or strong low-pass filtering and stimulus masked or not masked. Each subexperiment consisted of a 6 (stimulus exposure durations) × 3 (intact, LSF, or HSF) design. Each observer had at least 4 practice sessions in each subexperiment before beginning it. Each observer participated in the subexperiments in the same order: unmasked–mild low-pass filtering, unmasked–strong low-pass filtering, masked–mild low-pass filtering, and masked–strong low-pass filtering.

Each session consisted of 36 trials, comprising two instances of each of the 18 conditions. At the beginning of each session, trial ordering was randomized, the 36 digit strings were randomly generated, and the appropriate version of the stimulus for each trial—intact, HSF, or LSF—was created.

Procedure. Each trial consisted of a single stimulus presentation and test. A trial began with a 250-ms fixation cross accompanied by a 1,000-Hz

Table 1
Experiment 1: Various Experimental Parameters, Data, and Theory Fits

Variable	Unmasked Mild	Unmasked Strong	Masked Mild	Masked Strong
Observer EU				
Stimulus contrast	.058	.058	.058	.058
MSE	0.032	0.030	0.030	0.029
τ	6.2	8.7	7.9	9.2
θ	0.018	0.021	0.023	0.024
w_L	.563	.433	.537	.451
c	1.432	0.795	1.128	0.667
Y	1.000	.934	.969	.879
RMSE	0.0475	0.0348	0.0273	0.0336
F	2.25	1.31	0.81	1.39
Theory–data r^2	.985	.994	.994	.993
Observer JM				
Stimulus contrast	.073	.073	.073	.088
MSE	0.030	0.032	0.030	0.030
τ	3.4	3.3	7.1	6.0
θ	0.025	0.029	0.024	0.038
w_L	.610	.504	.647	.509
c	1.502	1.849	2.327	1.543
Y	.947	.961	.883	.949
RMSE	0.0605	0.0643	0.0389	0.0334
F	4.17	4.09	1.63	1.27
Theory–data r^2	.978	.967	.986	.992
Observer JF				
Stimulus contrast	.088	.104	.088	.088
MSE	0.031	0.031	0.031	0.027
τ	5.5	3.7	8.2	8.7
θ	0.029	0.041	0.050	0.050
w_L	.657	.564	.837	.657
c	2.478	2.491	2.853	2.211
Y	.739	.868	.838	.700
RMSE	0.0346	0.0401	0.0357	0.0226
F	1.28	1.68	1.36	0.69
Theory–data r^2	.988	.986	.979	.986
Mean data				
MSE	0.0177	0.0179	0.0176	0.0164
τ	5.0	5.2	7.7	8.0
θ	0.024	0.030	0.032	0.037
w_L	.610	.500	.673	.539
c	1.804	1.712	2.103	1.474
Y	.895	.921	.897	.843
RMSE	0.0356	0.0391	0.0143	0.0127
F	4.06	4.75	0.66	0.59
Theory–data r^2	.990	.988	.998	.999

Note. RMSE = root-mean-square error.

warning tone. Then, 500 ms after cross and tone cessation, the appropriate version of the stimulus appeared for its appropriate duration. In the mask subexperiments, a noise mask consisting of a random jumble of black lines on a white background then appeared for 500 ms, whereas in the unmask subexperiments, the blank field returned for 500 ms. The observer then attempted to type in the four just-seen digits in correct order. The observer was required to type in exactly four digits and was able to backspace and correct if desired. Following the observer's response was feedback in the form of a string of four Xs and Os; an O indicated that the digit in the corresponding position was correct, and an X indicated that it was incorrect.

Results and Discussion

We computed proportion digits recalled in their correct positions, corrected for the .1 guessing rate. For each subexperiment, we generated three *performance curves* (performance as a function of stimulus duration), one for each of the three stimulus types, for each observer. There were no systematic differences among observers, so the data were averaged over them. Each panel of Figure 3 shows the resulting data for one of the four subexperiments. Standard errors, averaged over the 18 conditions of each subexperiment, are shown in Table 1, Row 2, for each observer and for the mean data.

SRIA theory fit. As described, the SRIA theory incorporates four free parameters: τ, one of the impulse-response function parameters (the other, n, was set at 9); θ, the sensory threshold; c, the proportionality constant relating $\{A_\theta(t) \times [1 - I(t)]\}$ to $r(t)$; and w_L, the LSF weight. In addition, we needed one additional parameter, an *asymptote*, which we designate Y. Our four-digit strings could, with sufficient effort, be reported perfectly at long durations. However, because of attention lapses, motor errors, and the like, asymptotic performance was below perfect.

The best fitting parameters, along with three theory-fit measures, are shown in Table 1 for the 3 observers. The theory-fit measures are the following: Root-mean-square error (RMSE) is the square root of the mean theory–data squared deviations. F is the F value corresponding to the deviation of the data from the null hypothesis corresponding to the theory and is the squared ratio of the RMSE to the standard error. The theory–data r^2 is the across-conditions Pearson r^2 between the data points and the theoretical predictions. To arrive at the theory shown in Figure 3, we averaged the theoretical predictions across observers, thereby corresponding to the similarly averaged data. The mean theory-fit measures in Table 1 are based on comparisons between the mean theory and mean data.

To evaluate the theory fit, we focus on the mean data. The F values for the two unmasked conditions are both significant, which allows us to statistically reject the SRIA theory as not being sufficient to describe these data points. However, a glance at the Figure-3 data fits along with the barely visible standard errors makes it evident that we can reject the theory only because the data have enormous statistical power. An alternative fit measure that we consider more illuminating, the Pearson r^2s between the data and the theoretical predictions, ranges from .988 to .999. Thus, the theory captures a very large proportion of between-conditions variance.

Implications for spatial-frequency additivity. Experiment 1 indicates that the SRIA theory accounts satisfactorily for the relations among intact, HSF, and LSF performance. Of most importance is that the SRIA theory is an independence theory: It assumes (a) time-independent acquisition of HSF and LSF information and (b) HSF and LSF additivity, in that the sensory response function for the HSF + LSF sum (i.e., the intact stimuli) is the sum of the individual HSF and LSF sensory response functions. We have thus replicated empirical and theoretical findings reported by Olds and Engel (1998), and we have generalized them in three ways: first, to digit strings as well as pictures of objects; second, to unmasked as well as masked stimuli; and third, to differing degrees of low-pass filtering. This version of the SRIA theory is inconsistent with global-precedence GtL theories which assume that HSF informa-

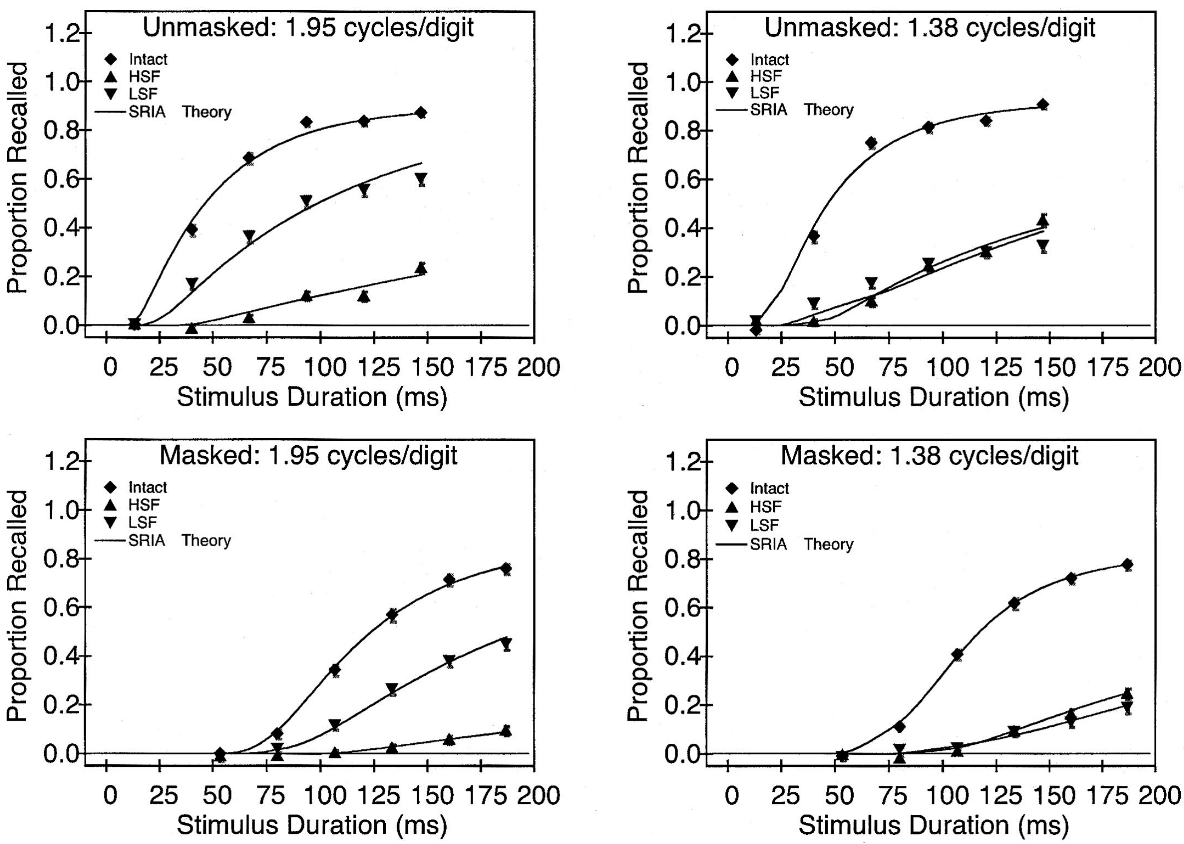

Figure 3. Experiment-1 data: Each panel shows performance, averaged over 3 observers, as a function of stimulus duration for three spatial-frequency conditions. Solid lines show best average fit of the sensory response–information acquisition (SRIA) theory. Left panels show data from mild low-pass filtering; right panels show data from greater low-pass filtering. HSF = high spatial frequency; LSF = low spatial frequency. Error bars are standard errors.

tion and LSF information are acquired over different time courses and, ipso facto, with stronger, interactive GtL theories which assume that initial acquisition of LSF information influences subsequently acquired HSF information. This version of the theory could not account for Schyns and Oliva's (1994) data described above, although it is sufficient to describe Olds and Engel's and the present data.

Experiment 2: Priming

Despite Olds and Engel's (1998) and our Experiment-1 data supporting an independence theory of spatial-frequency information acquisition, there is still reason to question such an account on several grounds. First, as described earlier, there is ample evidence in the literature for some form of GtL processing. In particular Sanocki (2001) provided strong evidence for interactive GtL theories using a paradigm in which a target picture (a simple line drawing of a house or a vehicle) was presented to an observer who was then required to distinguish the target from two same-shape distractors. Either just before or just after target presentation, there briefly appeared one of two kinds of prime. *Large-scale* primes depicted the global outline of the target, whereas *small-scale* primes depicted small interior details of the target. In the most compelling of Sanocki's (2001) three experiments, the large-scale prime provided no information that would allow the observer to distinguish the target from the distractor. Nevertheless, the large-scale prime, when presented before the target, improved performance. Sanocki concluded that the large-scale prime, when presented before the target, provided a perceptual framework within which the remainder of the target information could be interpreted.

One might argue that this finding is not directly relevant to the present work because Sanocki (2001) defined local and global somewhat differently than did Olds and Engel (1998), Schyns and Oliva (1994), and the present work—Sanocki used feature size rather than spatial frequency. However, the empirical work of Sanocki and others is not the only basis for reserving judgment about the implications of the Olds and Engel (1998) and of the present Experiment-1 results. Conclusions in both these instances were based primarily on the adequate fit of a quantitative theory; but an adequate fit of a quantitative theory constitutes only weak support for the theory's assumptions (see Roberts & Pashler, 2000). Perhaps in both Olds and Engel's experiments and in our Experiment 1, the flexibility of the SRIA theory's five free parameters allowed the theory to fit data issuing from what is still fundamentally some form of GtL processing.

As we describe, Experiment 2 was designed explicitly to test a prediction of interactive GtL theories. Given the relations among theories that we have described, confirmation of an interactive GtL theory would constitute confirmation of a global-precedence theory; however, disconfirmation of an interactive GtL theory would not rule out a global-precedence theory. It had been our intention, had we disconfirmed an interactive theory, to carry out further experiments to distinguish between independence theories and global-precedence theories. However, a serendipitous finding in Experiment 2 rendered this step unnecessary: To anticipate our conclusion, the Experiment-2 data allow us to disconfirm both independence theories and interactive theories, thereby leaving global-precedence theories, of which we provide an explicit, quantitative example, as the only viable explanation of both our Experiment-1 and Experiment-2 data, along with the Schyns and Oliva (1994) and the Olds and Engel (1998) data described earlier.

In Experiment 2, we investigated temporal instead of spatial interactions between HSF and LSF information using a priming paradigm related to one introduced by Parker et al. (1996). To do so, we incorporated four conditions. In two *no-prime* conditions, either HSF-only or LSF-only stimuli were, exactly as in Experiment 1, displayed for varying durations. In addition there were two *priming* conditions. In the first, a 40-ms LSF prime was followed immediately by a variable-duration HSF target. In the second priming condition, a 40-ms HSF prime was followed immediately by a variable-duration LSF target. The target was always exactly the same digit string as the prime, that is, the target and prime on a given primed trial differed only in spatial-frequency composition. No masks were used in Experiment 2.

These conditions allow testing of a strong and unambiguous interactive GtL theory prediction that uses methodology and associated logic introduced by Loftus, Johnson, and Shimamura (1985; see also Loftus, Duncan, & Gehrig, 1992) for investigating the contribution to memory of an iconic image. Loftus et al. (1985) presented stimuli at varying durations that were followed either by an immediate mask (which did not allow an icon) or by a 300-ms delayed mask (which did allow an icon). They then observed that performance curves for the immediate-mask and delayed-mask conditions were horizontally parallel, with the immediate-mask curve shifted to the right by 100 ms. That is, performance for a d-ms stimulus followed by an icon was identical to performance to a (d + 100)-ms stimulus not followed by an icon. Loftus et al. (1985) concluded that irrespective of the duration of the physical stimulus that it follows, “an icon is worth 100 ms of additional physical exposure duration” (p. 1).

Reinitz, Wright, and Loftus (1989) used similar logic to investigate priming effects. They presented pictures (e.g., of a guitar), at varying exposure durations, preceded by a prime (in this example, the word “GUITAR”) or not preceded by a prime. Reinitz et al.’s goal was to test two possible accounts of the prime’s effect on stimulus processing. The first was that the prime acted like a brief preview of the stimulus itself, whereas the second was that the prime acted to speed up stimulus processing. The test again involved a horizontal comparison of the primed and the unprimed performance curves. The rationale was that if the prime acted like a brief preview of the stimulus (akin to an icon’s acting as a brief postview), then the primed and unprimed performance curves would be horizontally parallel, separated by the amount of preview that the prime was worth. If the prime sped up processing, however, the curves would be horizontally diverging, with the (horizontally compared) slope of the primed curve being greater than the slope of the unprimed curve. Reinitz et al. found that the prime acted to speed up processing, thereby confirming an interactive theory.

Our present Experiment 2 used similar logic. To understand it, it is useful to begin with a thought experiment. Consider one of the two types of primed target stimuli, say, the HSF targets. Suppose that the prime that preceded the d-ms HSF target were simply a 40-ms HSF stimulus rather than the 40-ms LSF stimulus that we actually used. Now suppose we plot two performance curves: one corresponding to the primed targets and the other corresponding to the unprimed targets. Because the prime and the stimulus are identical and because the stimulus immediately follows the prime, it is obvious that the primed stimulus is identical to the unprimed stimulus, except for being 40 ms longer. Therefore, within the limits of statistical error, the two curves would be identical but separated horizontally by 40 ms.

Now consider our actual priming condition in which the HSF stimulus is preceded by a LSF prime. Suppose that the LSF prime did not influence HSF stimulus processing. In that case, the primed and unprimed curves would still be separated by some constant corresponding to how much HSF information a 40-ms LSF prime was worth, that is, the LSF prime would simply add to the HSF stimulus. If, on the other hand, the LSF prime influenced the processing speed of the HSF stimulus, then the curves would not be horizontally parallel. For example, if the LSF prime increased the processing rate of the HSF stimulus, the primed and unprimed curves would diverge horizontally.

Method

Observers. Observers were 6 paid University of Washington undergraduates and graduates with normal or corrected-to-normal vision.

Stimuli and apparatus. HSF and LSF stimuli were created as in Experiment 1. The low-pass filter size was 1.59 c/d. For each observer except JRS, the prime and stimulus had the same contrast, which is shown for individual observers in Table 2, Row 1. JRS’s prime contrast was accidentally set at .073 rather than to his target contrast of .088, but because of the theory-based way in which we analyzed our data, this error did not present any interpretational difficulties. The apparatus was the same as in Experiment 1.

Procedure. Each observer participated in 50 data-collection sessions. Each session included 48 trials: 2 in each of the 24 conditions defined by 4 priming conditions $\times$ 6 durations. The order of the 24 conditions was randomized for each session.

Each trial consisted of a presentation and test of a single letter string. A trial began with a 250-ms fixation cross accompanied by a 1,000-Hz warning tone. Then, 500 ms after cross and tone cessation, the letter string was presented in one of its various Prime Target $\times$ Duration configurations.

On an unprimed trial, the target letter string only—either a LSF or a HSF version—was presented at one of six durations. On a primed trial, a prime was presented first, always for 40 ms, followed immediately by the target, presented at one of six durations. The target was always the same letter string as the prime, but in the opposite spatial-frequency condition; that is, LSF primes were always followed by HSF targets and vice versa.

Because we wanted to map out approximately equal performance ranges in the primed and unprimed conditions, target exposure durations were somewhat longer in the unprimed compared with the primed conditions: There were six target durations that ranged from 27 to 160 ms in the unprimed conditions and from 0 to 133 ms in the primed conditions.

Table 2
Experiment 2: Various Experimental Parameters, Data, and Theory Fits

Variable	JRS	CAW	SKC	HYC	JEP	TMB	Mean data
	Observer						
Stimulus contrast	.088	.088	.088	.088	.119	.104	
MSE	0.0270	0.0259	0.0243	0.0236	0.0245	0.0245	0.0102
Curve shifting							
HSF: Best Δt (LSF prime's worth)	95	24	49	34	49	59	49
LSF: Best Δt (HSF prime's worth)	33	48	32	48	40	26	42
HSF: r^2	.990	.980	.991	.993	.939	.993	.993
LSF: r^2	.978	.970	.980	.922	.987	.977	.979
Interaction magnitude							
Observed interaction contrast magnitude	0.035	0.486	0.142	0.325	0.175	0.187	0.225
Crossover interaction?	No	Yes	Yes	Yes	Yes	Yes	Yes
SRIA theory interaction magnitude	−0.086	0.252	−0.005	0.033	0.072	0.038	0.051
Crossover interaction?	No	No	No	No	No	No	Yes[a]
Modified theory interaction magnitude	0.057	0.335	0.082	0.184	0.149	0.134	0.157
Crossover interaction?	No	Yes	Yes	Yes	Yes	Yes	Yes
Original SRIA theory fit							
RMSE (data–theory)	0.0719	0.0470	0.0413	0.0620	0.0645	0.0494	0.0365
F (theory)	49.69	21.19	16.37	36.99	39.94	23.49	76.89
r^2 (data–theory)	.904	.971	.977	.954	.920	.972	.979
Modified SRIA theory fit							
RMSE (data–theory)	0.0468	0.0452	0.0287	0.0386	0.0447	0.0408	0.0229
F (theory)	3.01	3.04	1.39	2.67	3.32	2.77	5.07
r^2 (data–theory)	.959	.973	.989	.982	.962	.981	.992
Modified theory parameter values							
τ	11.5	20.1	22.4	12.4	11.9	18.5	16.1
θ	0.0055	0.0084	0.0064	0.0104	0.0183	0.0084	0.0096
$w_L(100)$	.724	.653	.750	.540	.611	.753	.667
c	2.78	3.84	1.87	1.49	3.22	1.55	2.46
Y	.794	.999	.812	.902	.743	.877	.854

Note. HSF = high spatial frequency; LSF = low spatial frequency; SRIA = sensory response–information acquisition; RMSE = root-mean-square error.
[a] The crossover interaction for the mean data was due to an averaging artifact.

To forestall confusion, we emphasize a slightly odd implication of this timing arrangement that involves the two zero-target-duration primed conditions. In these two conditions, *zero duration* means that the target does not appear at all but rather the prime alone is presented. The potential source of confusion is that the zero-duration-primed HSF condition involves *only* a 40-ms LSF stimulus and, likewise, the zero-duration-primed LSF condition involves only a 40-ms HSF stimulus.

Results and Discussion

Figure 4 shows Experiment-2 performance curves for the four priming conditions averaged over the 6 observers, along with theory fits that we describe shortly. Standard errors, averaged over the 24 conditions, are shown in Table 2, Row 2, for each observer and for the mean data. Primed conditions are shown in Figure 4 by solid-triangle curve symbols (upward-pointing for HSF conditions and downward pointing for LSF conditions), whereas unprimed conditions are shown by analogous open-triangle curve symbols. Note that, as described above, the two zero-duration primed conditions (represented by the two solid triangles at the far left of Figure 4) consist of the prime only: The 0-ms HSF primed condition (filled, upward-facing triangle) is actually a single, 40-ms presentation of a LSF stimulus, and the 0-ms LSF primed condition (filled, downward-facing triangle) is a single, 40-ms presentation of a HSF stimulus.

Parallel primed and unprimed curves. Recall that the prediction of a noninteractive theory—that is, an independence or a global-precedence theory—is that the primed and unprimed curves be horizontally parallel for both HSF and LSF stimuli. Are they? A casual inspection of Figure 4 suggests that they are. In quest of a quantitative answer, we did the following for each observer.

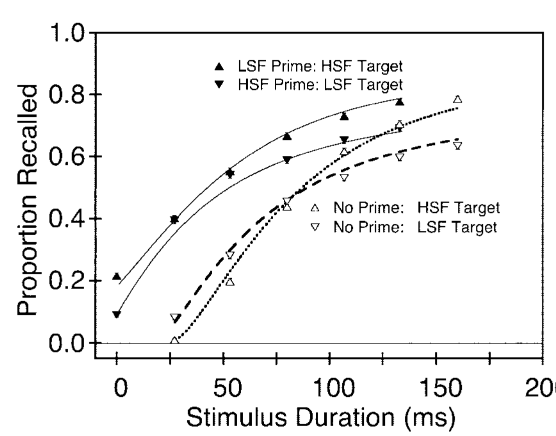

Figure 4. Experiment-2 data: Performance, averaged over 6 observers, as a function of stimulus duration for four priming conditions. Solid, dashed, and dotted lines show best average fit of the modified sensory response–information acquisition theory. HSF = high spatial frequency; LSF = low spatial frequency. Error bars are standard errors.

1. Consider, to illustrate, the HSF performance curves. There were six stimulus-presentation durations for the unprimed stimuli, which can be thought of as a six-element vector d_{UP}, whose entries range from 27 to 160 ms. Similarly, the six durations for the primed stimuli form a six-element vector, d_P, whose entries vary from 0 to 133 ms. If the primed and unprimed HSF curves are horizontally parallel, then there should be some value, Δt, so that if the primed curve is shifted by Δt to the right, it will overlap perfectly with the unprimed curve; that is, the 12 data points—the 6 unprimed data points plotted as a function of d_{UP} and the primed data points plotted as a function of $(d_P + \Delta t)$—will form a single monotonic function. The identical logic, of course, holds for LSF stimuli.

2. Past experience (see, e.g., Busey & Loftus, 1994) indicates that these performance curves can be fit well by the linear function, $-\ln(1 - p/Y) = kd$, where p is performance, k is a constant, d is stimulus duration, and Y is asymptotic performance. Thus, for any shift, Δt, we can fit this linear function to the 12 data points—the 6 unprimed and 6 primed $-\ln(1 - p/Y)$ data points plotted simultaneously as functions of d_{UP} and $(d_P + \Delta t)$, respectively.

3. Accordingly, for each observer, the best fitting values of k, Y, and Δt—those that maximized the Pearson r^2 between the combined durations and the combined data points—were found for both the HSF and LSF curves. The resulting Δt and r^2 values are shown in Table 2 in the section labeled "Curve shifting."

4. Finally, for each spatial frequency, we plotted mean unprimed curves against the original unprimed duration values, along with the mean primed curves against the original primed durations shifted rightward by the mean Δd. These curves are shown in Figure 5 (to avoid visual clutter, the HSF curves are shifted upward by 0.5).

It is evident both from visual inspection of the Figure-5 curves and the Table-2 curve-shifting r^2 values (both for individual observers and for the mean data) that for both spatial frequencies, the alignment is good. For all intents and purposes, the primed and unprimed curves are horizontally parallel for both spatial frequencies. This finding appears to disconfirm the interactive GtL prediction described above.

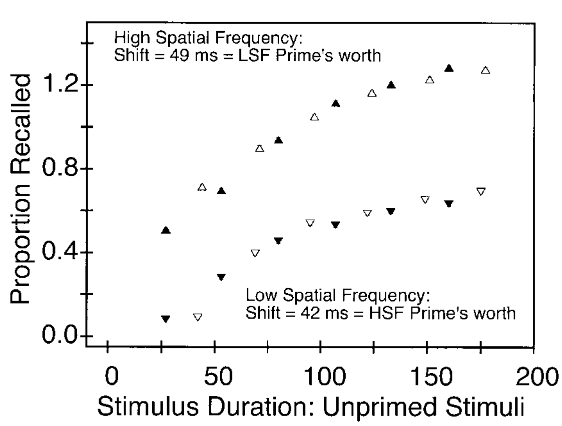

Figure 5. Experiment-2 data: curve-shifting results. For each spatial frequency, the primed curves have been shifted rightward to provide the best fit with the unprimed curves. Error bars are standard errors. For visual clarity, the high-spatial-frequency (HSF) curves were raised above the low-spatial-frequency (LSF) curves by 0.5.

Before unequivocally accepting this conclusion however, we consider one possible challenge to it, involving the notion of prime-target masking: Perhaps the HSF primes forward-masked the LSF targets and/or vice versa, and the resulting degradation of primed compared with unprimed target information somehow resulted in the parallel curves.

Our response to such a challenge would be as follows. We have formulated a very strong prediction of a noninteractive theory, that of horizontally parallel curves. As Figure 5 shows, this prediction was resoundingly confirmed. In general, confirmation of strong predictions affords strong support to the theoretical logic underlying the prediction. In contrast, no masking account that makes this prediction readily presents itself. If one were to advocate an alternative, say, masking account of our results, it would be the advocate's responsibility to formulate the alternative account in such a way that it made the same specific prediction. Moreover, any masking account would be relatively complex, in that it would have to posit two opposing influences—the positive effect of prime qua prime along with the negative effect of the prime as mask—that would somehow combine to imply parallel curves. In view of these considerations, the present theoretical logic—that the parallel curves issue simply and naturally from a noninteractive process—seems most parsimonious and most viable.

Spatial-Frequency × Duration interactions. As indicated, our finding of horizontally parallel primed and unprimed curves, while ruling out an interactive GtL theory, cannot be used to distinguish between independence theories and global-precedence GtL theories. However, a serendipitous result issuing from the two unprimed conditions allows us to make this distinction. In particular, as can be seen in Figure 4, there is a crossover interaction between spatial frequency and duration: At short durations, LSF performance exceeds HSF performance and vice versa at longer durations. Such an interaction is inconsistent with independence theories and, accordingly, warrants careful scrutiny, particularly because a crossover interaction can emerge as an averaging artifact.

The first row of the Table-2 "Interaction magnitude" section indicates this crossover interaction to have occurred for 5 of the 6 observers. To quantify it, we applied the 1-degree-of-freedom interaction contrast $(-1, 1, 1, -1)$ to the conditions: HSF–lowest duration, LSF–lowest duration, HSF–highest duration, and LSF–highest duration. The contrast magnitude is shown for the 6 observers in Table 2: It is positive even for JRS, the one observer who did not show the interaction in crossover form. The mean interaction magnitude is 0.225 with a 95% confidence interval of ± 0.127. Thus, the interaction is generally crossover and consistent across observers. An intuitive interpretation of this result is that it is more efficient to acquire information from LSFs at short durations but more efficient to acquire information from HSFs at longer durations. This finding dovetails with that described above, reported by Schyns and Oliva (1994): The message is that HSFs and LSFs are differentially effective at different times after stimulus onset.

SRIA theory fit. Like any independence theory, the SRIA theory used by Olds and Engel (1998) cannot fit the just-described interaction: If the weighting used for the LSF and HSF components remains constant, then whichever component has the higher weight must produce higher performance for all durations. To demonstrate this, we fit the theory to the data for each of the 6

Experiment-2 observers. We assumed that (a) the prime and the stimulus triggered sensory response functions in the usual manner, (b) in the primed conditions, the overall sensory-response function was the appropriately weighted sum of the prime and stimulus sensory response functions, and (c) performance was determined by above-threshold area under the overall sensory response function.

The SRIA theory fit measures shown in the Table-2 "Original SRIA theory fit" section are not as good as those of Experiment 1, as indicated by all fit measures. This is not surprising in light of the just-described crossover interaction that the SRIA theory, an independence theory, is incapable of predicting. The SRIA theory's failure to predict the crossover interaction is shown in Table 2, second part of the "Interaction magnitude" section. Note that the failure occurs for all observers individually, although an averaging artifact allows a (very small) crossover interaction for the mean predictions.

A Modified SRIA Theory

The poor fit of the SRIA theory to Experiment 2 in general and its failure to allow the observed crossover interaction in particular indicate that the process generating the Experiment-2 data is one in which information based on different spatial frequencies have different time courses—i.e., that a global-precedence GtL theory is appropriate to describe the data. To illustrate this conclusion's validity, we modified the SRIA theory to incorporate such a property. Instead of assuming simple weightings, w_L and $(1 - w_L)$ for the LSFs and HSFs, we allowed the weight to vary over time t since stimulus onset. In particular, we assumed that the weight for the LSF sensory-response function at time t was $w_L(t) = e^{-kt}$, whereas the HSF weight was $w_H(t) = [1 - w_L(t)]$. For ease of exposition, we reparameterized the exponential defining w_L: Instead of using a decay rate, k, we used what we term $w_L(100)$, which is defined to be the LSF weight 100 ms after stimulus onset.

Note that this modification does not increase the number of free parameters; it merely substitutes $w_L(100)$ for w_L. Note also that the original and modified theories are mutually exclusive. The modified theory cannot produce the same state of the system as the original theory, except for the degenerate cases of $w_L(100) = 0$ and $w_L(100) = 1$. This is as it should be: Because the original theory is an independence theory and the modified theory is a GtL theory, the two must be mutually exclusive.

Application of the Modified SRIA Theory to Experiment 2

Application of the modified SRIA theory to the Experiment-2 data is straightforward: It is as described above; the only difference is the weighting scheme. The predictions are shown as the solid, dashed, and dotted lines in Figure 4, and various fit measures are provided in Table 2. The third part of the Table-2 "Interaction magnitude" section demonstrates that the observed crossover interaction is tracked perfectly by the modified theory across the 6 observers. The "Modified SRIA theory fit" and "Modified theory parameter values" sections provide fit measures and parameter values. All in all, the modified theory fit is better than the original SRIA theory fit for all observers and is quite acceptable in general by all fit measures.

Application of the Modified SRIA Theory to Other Data

A modified theory designed to fit the Experiment-2 data is not satisfactory if it fits the Experiment-2 data but no other data. Accordingly, we fit the modified theory to four other data sets: the present Experiment 1 and Olds and Engel's (1998) Experiments 1–3. The magnitudes of the fits, expressed as r^2s, are provided in Table 3, and we describe them in turn.

Present Experiment 1

We fit the modified theory to the three Experiment-1 observers and the mean modified theory to the mean data. The r^2s are averaged over the four Mask $\times$ LSF Strength conditions. As indicated in Table 3, Rows 1–4, the fits of the two theories are very similar. The mean fit of the modified theory is slightly better than the mean fit of the original theory. The estimates of the four parameters common to the two theories—τ, θ, c, and Y—were quite similar when estimated in the context of the modified theory to their estimates using the original theory. The estimates of $w_L(100)$ were .636, .526, .725, and .603 for the unmasked (mild low-pass filtering), unmasked (strong low-pass filtering), masked mild, and masked strong subexperiments. These are comparable in magnitude to the corresponding estimates of $w_L(100)$ obtained in Experiment 2 (see Table 1).

Olds and Engel's (1998) Experiment 1

Earlier, we described Olds and Engel's (1998) Experiment 1: Essentially, it was the same as our Experiment 1 with masked line drawings as stimuli and with a single degree of LSF filtering strength. As indicated in Table 3, Rows 5–9, the fits of the original and modified theories are very similar, and the mean fit of the modified theory is virtually identical to the mean fit of the original theory.

Olds and Engel's (1998) Experiment 2

Olds and Engel's (1998) Experiment 2 was similar to their Experiment 1 except that they used 48 naive observers instead of

Table 3
Comparisons of Original and Modified Theories for Various Data Sets (r^2 Values)

Experiment	Original theory	Modified theory
Experiment 1		
Obs EU	.992	.986
Obs JF	.985	.986
Obs JM	.981	.986
Mean	.979	.992
O&E Experiment 1		
Obs EO	.997	.997
Obs SE	.992	.988
Obs KS	.964	.960
Obs PB	.997	.993
Mean	.996	.995
O&E Experiment 2	.970	.965
O&E Experiment 3		
Obs SE	.950	.931
Obs CF	.923	.910
Mean	.950	.932

Note. Obs = observer; O&E = Olds & Engel (1998).

4 practiced observers. As indicated in Table 3, Row 10, the fit of the original theory is slightly better than the fit of the modified theory.

Olds and Engel's (1998) Experiment 3

In their Experiment 3, Olds and Engel (1998) used 2 practiced observers and varied the relative amounts of HSF and LSF contrast. In particular, they created HSF and LSF components of their intact stimuli and then created eight versions in which they added the HSF components scaled by 0, 0.07, or 0.14 to the LSF component also scaled by 0, 0.07, or 0.14 (the eight versions resulted from the 3 $\times$ 3 [HSF Contrast $\times$ LSF Contrast] combinations absent the no-stimulus-at-all version in which both components are scaled by 0). Each of these eight versions was shown at eight exposure durations ranging from 13 to 426 ms, yielding 64 conditions in all. As indicated in Table 3, Rows 11–13, the fits of the original theory are slightly better than those of the modified theory. Figure 6 shows the fit of the data for the 2 observers. It is generally satisfactory—approximately as good as that reported by Olds and Engel for the original theory—although with some notable systematic exceptions at low-contrast levels.

These exceptions, however, are intriguing. Consider first the lower two Figure-6 panels wherein the HSF component is 0, that is, there is only an LSF component. Across the 2 observers, three out of the four curves asymptote at considerably less-than-perfect performance levels. This cannot happen within the context of the original SRIA theory (unless an asymptote is explicitly built in, as it was in the fit to the present experiments, in which case, all conditions will asymptote at this same level). In the present discussion, we, for ease of exposition, assume there to be no built-in asymptote. As long as stimulus contrast is above the threshold value, θ, longer durations lead to longer above-threshold areas, and performance must approach 1.0. With the modified SRIA theory, however, the weight accorded the LSF component declines over time since stimulus onset; therefore, as duration becomes longer, the corresponding sensory response functions begin to decline. As they decline below threshold, above-threshold area, and therefore performance, no longer increases. The result is a below-perfect asymptote, as shown by the theoretical predictions in Figure 6, lower panels.

Alas, though, the story is not this simple: There also appears to be a below-perfect asymptote for two out of the four curves involving a zero-LSF component: As indicated in Figure 6, middle panels, for both observers, the zero-LSF, 7% HSF condition also appears to asymptote below 1.0. This is not permitted by the modified SRIA theory, as indicated by the rising-to-1.0 theoretical predictions. We return to this issue shortly.

Modified Theory Fits: Discussion

The modified theory fits our Experiment-2 data well. It also fits Olds and Engel's (1998) Experiments 1 and 2 quite well. These good fits are somewhat surprising. In devising our modified theory, we set out to find a general class of theories that would at least qualitatively explain both the crossover interaction and the parallel primed–unprimed curves that we observed in our Experiment 2—in other words, a theory that would constitute a global-precedence GtL theory rather than an independence theory. Within the context of the SRIA theory, this required only that the relative weightings of HSF and LSF components of the sensory response function change over time. There is an infinite number of quantitative incarnations that would satisfy this property, and we chose the particular one that we did—exponential decay of the LSF weights—because it is a convenient one-parameter function. That the quantitative fit to Experiment 2 was so good should not, therefore, be attributed to particularly insightful theory construction on our part but rather to a stroke of good luck.

Despite our original conclusions about the Olds and Engel (1998) data and our Experiment-1 data—that they are describable with an independence theory—we wish to emphasize that we do not consider our Experiment-2 results to be in any way at odds with our Experiment-1 results or with Olds and Engel's results. Our modified theory accounted well for all data, and in general, it is not difficult to devise a data set that, while generated by one theory, is describable by another. Because the modified theory is sufficient to describe all data, while the original theory cannot describe the Experiment-2 data, we conclude that the modified theory is most viable.

The modified theory's major failing—both qualitatively and quantitatively—is in its fit to Olds and Engel's (1998) Experiment 3. As sketched above, although the modified theory qualitatively, although not quantitatively, accounts for the below-perfect asymptote of the LSF-only data, it cannot account for the similar below-perfect asymptote of the HSF-only data. As we have indicated, the modified SRIA theory accounts, at least qualitatively, for the LSF-only data (Figure 6, bottom panels) because the modified theory predicts a less-than-zero asymptote. The theory still fails though in that it cannot predict a less-than zero asymptote for HSF-only data, and such an asymptote clearly materializes in the two 7% HSF, zero-LSF curves in the middle panels.

Why is this? The only explanation that occurs to us is that despite being created to be homogeneous, the stimuli used by Olds and Engel (1998) vary slightly in their contrast levels so that some of them remain below threshold throughout a presentation, whereas others rise above threshold. The result would be average data with a below-perfect asymptote. This explanation, of course, applies to the LSF-only data as well as to the HSF-only data, thereby rendering unnecessary, at least as a qualitative explanation of these specific data, the HSF–LSF asymmetry inherent in the modified SRIA theory. Nevertheless, the modified SRIA theory fits the remainder of the data sufficiently well that extensions of Olds and Engel's Experiment 3 may show it to be necessary.

General Discussion

We began by dividing theories of how spatial-frequency components combine and contribute to visual information acquisition into two mutually exclusive categories: independence theories and GtL theories. We further divided GtL theories into global-precedence theories and interactive theories. We have concluded, in concert with others (e.g., Hughes et al., 1996; Schyns & Oliva, 1994), that a global-precedence theory coherently describes a wide range of visual perception data. In particular, in our Experiment 1, we replicated and extended parametric data reported by Olds and Engel (1998). Although Olds and Engel's and our Experiment-1 data are describable by an independence theory, our Experiment-2 data allowed us to disconfirm both an independence theory and an

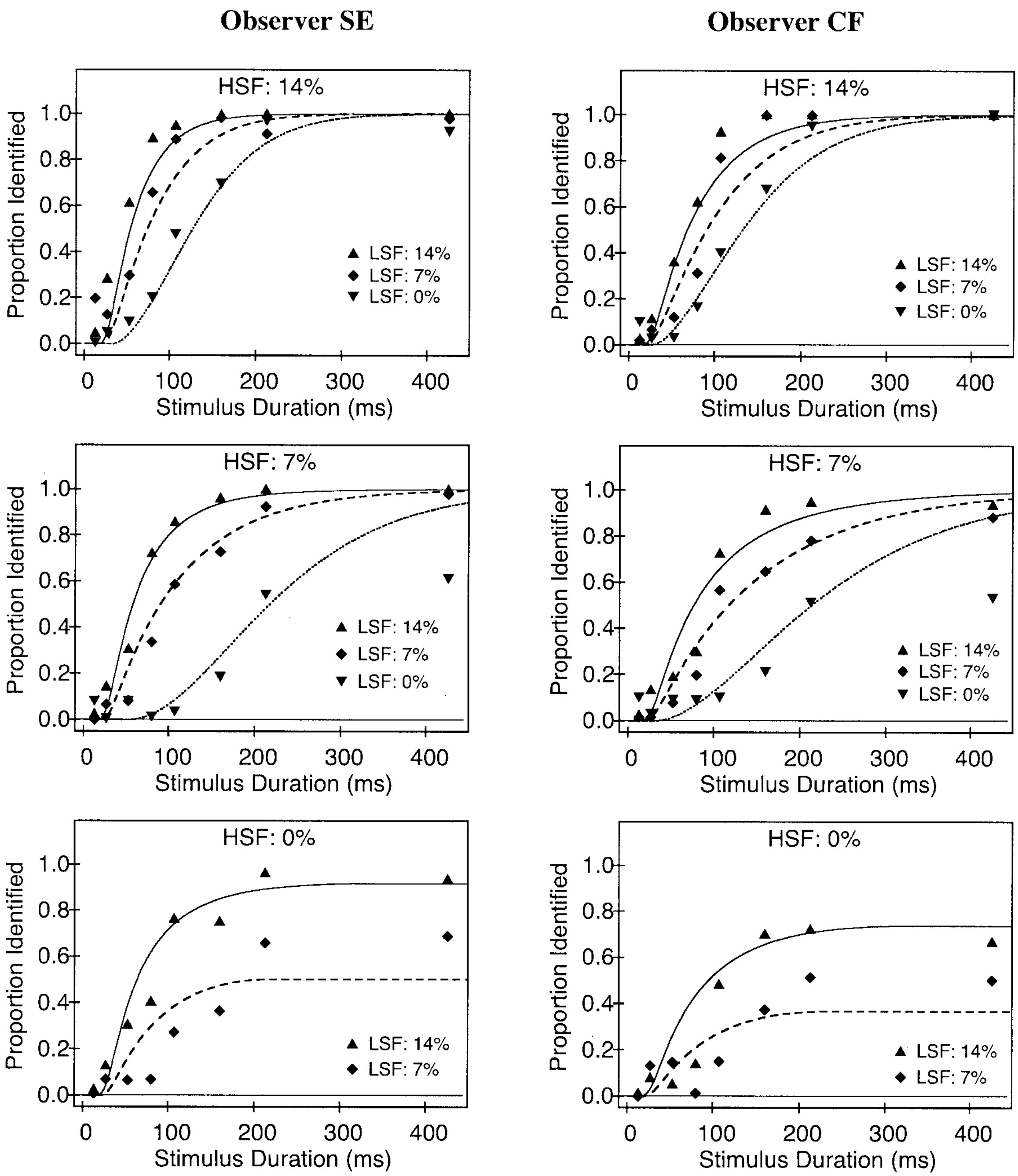

Figure 6. Data for 2 observers in Olds and Engel's (1998) Experiment 3. For each observer, each panel shows one level of high-spatial-frequency (HSF) contrast; within each panel, each curve represents one level of low-spatial-frequency (LSF) contrast. Smooth lines through the data points represent best fits of the modified sensory response–information acquisition theory.

interactive theory. It is possible that our data, involving digits, and Olds and Engel's data, involving pictures, operate according to different rules. However, we demonstrated that we could account for Olds and Engel's data as well as our own using a specific global-precedence theory; hence, as we indicated earlier, the most parsimonious account is that a global-precedence theory governs both data sets. As we indicated in our introduction, a global-precedence theory is also appropriate to describe the Schyns and Oliva (1994) data. But because Schyns and Oliva's experiment involves so few data points, we cannot apply our quantitative theory to it.

This may not be the end of the story. It is entirely conceivable that another, more complex experiment using this general paradigm would require an interactive GtL theory to describe it (as an interactive theory was necessary to describe the Sanocki, 2001, data) and that such a theory could be devised that would adequately account for all data sets. One could make the case that in response to Sanocki's data and his arguments (which we consider in more detail in the next section), one should develop such an interactive theory now. However in view of the quite substantial differences between our and Sanocki's paradigm, stimuli, task, and definition of global and local, we feel that such development is not yet obligatory.

The Case for Global–Local Interactions

In our introduction, we described interactive models of information acquisition. Sanocki (1993) has provided a compelling rationale for global–local interactions particularly for objects and scenes. Sanocki (1993) has reiterated the oft-noted fact that objects in the world can appear in an infinitude of orientations, sizes, shapes, colors, and so on and has underscored the obvious implication: "If during object identification, the perceptual system considered such factors for an unconstrained set of alternatives, the enormous number of combinations of stimulus features and feature-object mappings would create a combinatorial explosion" (p. 878). Sanocki has noted that an obvious means of reducing what would be an otherwise impossible information-processing task is to use early information to constrain the interpretation of later information. Sanocki has referred to this idea as the *contingency hypothesis* and has indicated that local-information acquisition being contingent on global-information acquisition seems to be the most reasonable interpretation of the hypothesis. In support of the contingency hypothesis, Sanocki (1993, 2001) has provided results from a priming paradigm in which a global prime (e.g., a sketchy outline of a house) preceded or did not precede a target that conformed to the global aspects of the prime but that contained additional, local details. The task was then to distinguish between two detailed alternatives—the target plus a distractor—whose global characteristics conformed equally well to the prime. He found that the prime facilitated such discrimination even though the prime in and of itself provided no information that would allow target–distractor discrimination. This occurred for both simple shapes (Sanocki, 1991) and for relatively complex objects such as vehicles and buildings (Sanocki, 1993, 2001). Sanocki concluded that his results provided evidence for an interactive theory: The presence of the prime could only have facilitated acquisition of local information from the target.

Others have come to the same general conclusion: For instance, Reinitz et al. (1989) found, as indicated earlier, that a verbal category prime ("GUITAR") facilitated later discrimination between two pictures of a guitar—the target plus a distractor guitar. And yet, data issuing both from objects (Olds & Engel, 1998) and from digits (present experiments) can be accounted for quite precisely without assuming that HSF information acquisition is affected by the presence of LSF information.

The most harmonious resolution of this apparent conflict is to assume that global information can facilitate acquisition of subsequent local information but either that the kind of global information which does the facilitating is not necessarily LSF information or that the kind of local information which benefits from such facilitation is not necessarily HSF information. As we indicated earlier, Sanocki (1993) hinted that spatial frequency was a candidate means of separating global from local information, noting in a discussion of informational "grain size" that "precise [perceptual] analyses would have a small grain size, whereas imprecise analyses would blur high-frequency details and have a relatively large grain size" (p. 882); however in Sanocki's actual experiments, global and local information were defined by outlines and detailed representations of target stimuli, not specifically high and low spatial frequencies.

Flexible Processing Orders

Up to this point, we have proceeded under the implicit (and common) assumption that visual information processing proceeds along a fixed route with respect to spatial scale. This assumption makes sense based on logic and a great deal of past data.

However, Schyns and his colleagues have recently argued and presented data favoring the proposition that the visual system does not necessarily operate so rigidly; instead, people may use spatial scales (e.g., as instantiated in different spatial-frequency bands) in different orders depending on the task they are trying to carry out (Gosselin & Schyns, 2001; Oliva & Schyns, 1997; Schyns & Oliva, 1999; see Morrison & Schyns, 2001, for a review).

Schyns and his colleagues have provided numerous demonstrations of such flexibility using faces as stimuli. For example, Schyns and Oliva (1999) showed observers *hybrid* faces: a double exposure of two superimposed faces, one composed of only LSFs and the other composed only of HSFs. The two faces differed on three dimensions: male–female, expressive–nonexpressive, and angry–happy. On each of a series of trials, observers categorized the faces along one of these dimensions, and the investigators noted whether the face corresponding to the chosen category was the LSF member or the HSF member of the hybrid. Their main finding was that the categorization task that was used influenced which member of the pair was chosen. For instance, in one of their experiments, observers chose the LSF face 38% of the time when performing an expression–no expression categorization but chose the very same LSF face 66% of the time when performing the happy–angry categorization. The investigators concluded that which image spatial-frequency band dominated observers' perception was strongly influenced by the task that they were carrying out.

Global-to-Local Versus Coarse-to-Fine

Morrison and Schyns (2001) have also considered the relation between two temporal processing shifts that have appeared in the literature: GtL processing and what they have characterized as coarse-to-fine (CtF) processing. They have pointed out that although GtL and CtF shifts may appear at first glance to simply be two names for the same thing, they may be quite different and possibly orthogonal to one another (p. 459). GtL processing, in their view, is a processing change that occurs over two-dimensional space: Whole parts of a scene are initially processed, followed by processing that is focused on increasingly smaller areas. CtF processing, in contrast, can occur in any region of space and may be appropriately realized in terms of a transition from processing LSFs to processing HSFs. By this view, the phenomena described and studied by Sanocki (e.g., 1993) would involve a GtL processing shift and would be describable by interactive theories.

In contrast, the phenomena described and studied by Schyns and Oliva (1994), Olds and Engel (1998), and the present work would involve a CtF processing shift and would be describable by global-precedence theories.

Theoretical Cacophony

The points made in the preceding two sections do not reflect a great deal of theoretical harmony with respect to the investigation of object perception insofar as it is affected by spatial scale. To summarize, there are (at least) three areas of current theoretical confusion that may be responsible for some of the seeming empirical incompatibilities and failures to replicate in the literature. The first is the GtL versus CtF distinction just discussed. The other two are as follows.

Definition of Spatial Scale

To study the kind of GtL (or CtF, or any other kind of processing that in some way incorporates the notion that the visual world is made up of different spatial scales), it is necessary to operationalize how different spatial scales are to be implemented. Numerous investigators have used spatial frequency as a means specifying different scales. Presumably, they have done so because spatial frequency is well defined, mathematically realizable, and likely directly related to the functioning of the visual system.

Nevertheless, there are other ways of operationalizing spatial scale. A common one is as implemented by Sanocki (1993) wherein, as described earlier, *large-scale* is defined by large, overall shapes of target stimuli, whereas *small-scale* is defined by small target details. Although such an approach has been useful, as demonstrated by Sanocki's contributions, it is not yet clear (a) how such an implementation can be quantified in a manner akin to the quantification of spatial-frequency analysis or (b) how results obtained with this representation of spatial scale relate to results obtained with other, for example, spatial-frequency representations.

Processing Flexibility

Even maintaining a consistent definition of spatial scale, viz., spatial frequency, the role of spatial scale in object perception is far from well defined, and may be not at all simple. The claims and associated demonstrations of Schyns and his colleagues that scale usage order may depend on the exact task being carried out mitigate against a theoretical strategy of looking for *the* way in which the visual system progresses through different spatial scales in performing any arbitrary task. Rather, they suggest that a more fruitful approach would be to determine exactly what information in the stimuli—defined in terms of spatial frequency or some other characterization of spatial scale—is most efficient for task performance and then, given such a determination, whether the observed order of scale usage is optimal.

References

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203,* 237–260.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

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 and Performance, 24,* 1188–1214.

Campbell, F., & Robson, J. (1968).Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551–566.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. *Annual Review of Psychology, 31,* 309–341.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision.* New York: Oxford University Press.

Gosselin, F., & Schyns, P. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271.

Graham, N. (1989). *Visual pattern analyzers.* New York: Oxford University Press.

Harley, E. M., Dillon, A. M., & Loftus, G. R. (2003). *Why is it difficult to see in the fog? How contrast affects visual perception and visual memory.* Manuscript submitted for publication.

Hughes, H. C., Nozawa, G., & Ketterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience, 8,* 197–230.

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 and Performance, 18,* 530–549.

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

Loftus, G. R., & McLean, J. E. (1999). A front end to a theory of picture recognition. *Psychonomic Bulletin & Review, 6,* 394–411.

Loftus, G. R., Nelson, W. W., & Kallman, H. J. (1983). Differential acquisition rates for different types of information from pictures. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 35*(A), 187–198.

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

Majaj, N. J., Pelli, D. G., Jurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research, 42,* 1165–1184.

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review, 8,* 434–469.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception, *Cognitive Psychology, 9,* 353–383.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Olds, E. S., & Engel, S. A. (1998). Linearity across spatial frequency in object recognition. *Vision Research, 38,* 2109–2118.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges: Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34,* 72–107.

Olzak, L. A. & Thomas, J. P. (1986). Seeing spatial patterns. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 7-1–7-56). New York: Wiley.

Parker, D. M., & Costen, N. P. (1999). One extreme or the other or perhaps the golden mean? Issues of spatial resolution in face perception. *Current Psychology, 19,* 118–127.

Parker, D. M., Lishman, J. R., & Hughes, J. (1996). Role of coarse and fine information in face and object processing. *Journal of Experimental Psychology*: *Human Perception and Performance, 22,* 1448–1466.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Reinitz, M. T., Wright, E., & Loftus, G. R. (1989). The effects of semantic priming on visual encoding of pictures. *Journal of Experimental Psychology: General, 118,* 280–297.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107,* 358–367.

Sanocki, T. (1991). Effects of early common features on form perception. *Perception & Psychophysics, 50,* 490–497.

Sanocki, T. (1993). Time course of object identification: Evidence for a global to local contingency. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 878–898.

Sanocki, T. (2001). Interaction of scale and time during object identification. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 290–302.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependent scene recognition. *Psychological Science, 5,* 195–200.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentation. *Cognition, 69,* 243–265.

Solomon, J. A., & Pelli, D. G. (1994, June 2). The visual filter mediating letter identification. *Nature, 369,* 395–397.

Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America, 4,* 2006–2021.

Appendix

Olds and Engel's Theory: Number of Free Parameters

Olds and Engel (1998) fixed n and τ to 9 and 3, respectively, and then fit the sensory response–information acquisition (SRIA) theory to their data with four parameters: w_H and w_L, the weights for high-spatial-frequency (HSF) and low-spatial-frequency (LSF) components; θ, the sensory threshold; and c, the scaling factor. Here we demonstrate two things: First that $w_H + w_L$ corresponds to intact stimulus contrast and, second, that Olds and Engel's version of the SRIA theory was a three-parameter, not a four-parameter, version.

Intact Stimulus Contrast Is $w_L + w_H$

In Olds and Engel's (1998) study, an intact stimulus was the sum of the HSF and LSF components. In their application of theory, the sensory response function for the intact component was

$$a(t)(\text{intact}) = w_H[f(t) * g(t)] + w_L[f(t) * g(t)] = (w_H + w_L)\,[f(t) * g(t)],$$

which is exactly our equation for an intact stimulus (see Equation 1 above), where $C = w_H + w_L$.

Number of Free Parameters

Suppose that a best fitting set of parameters is found; call them w_L, w_H, θ, and c. In this case, response probability for a stimulus shown at duration d will be,

$$p = 1 - e^{A_\theta(\infty)/c},$$

where

$$A_\theta(\infty) = \int_{t_1}^{t_2} a_\theta(t)dt = \int_{t_1}^{t_2} \{C[f(t) * g(t)] - \theta\}\, dt.$$

Here, $C = w_H + w_L$, t_1 is the time at which $a(t) = \theta$ as $a(t)$ rises, and t_2 is the time at which $a(t) = \theta$ as $a(t)$ decays; thus, for all ts such that $t_1 \le t \le t_2$, $a(t) > \theta$.

Now consider a new set of parameters, scaled by some positive constant, k: kw_L, kw_H, $k\theta$, and kc. Note first that $a'(t) = ka(t)$, where $a'(t)$ is the sensory-response function generated by the new parameter set. Note second that t_1 and t_2 will remain the same, because $a(t_1) = \theta$, $a'(t_1) = ka(t_1) = k\theta$, and likewise for t_2. Therefore,

$$A'_\theta(\infty) = \int_{t_1}^{t_2} \{kC[f(t) * g(t)] - k\theta\}dt$$

$$= k\int_{t_1}^{t_2} \{C[f(t) * g(t)] - \theta\}dt = kA_\theta(\infty),$$

and response probability is

$$p' = 1 - e^{kA_\theta(\infty)/(kc)} = 1 - e^{A_\theta(\infty)/(c)} = p.$$

Thus, scaling the parameters leaves response probability unchanged, which means that the best fitting free parameters are unique only up to a positive scaling factor.

If contrast, C, is known, then because $(w_H + w_L) = C$, w_L and w_H may be expressed as percentages of contrast, in which case $w_H = 1 = w_L$. If contrast is unknown, as it apparently was in Olds and Engel's (1998) experiment, it can simply be stipulated to be some constant, and the same argument holds. In either case, w_H and w_L, are no longer separate free parameters; rather, they are entirely dependent. The advantage of a known contrast is that the estimate of θ is in standardized units, viz., stimulus contrast.

Received March 5, 2003
Revision received July 24, 2003
Accepted August 1, 2003 ■