Cued detection with compound integration-interruption masks reveals multiple attentional mechanisms

Philip L. Smith

The University of Melbourne, Australia

Rachel Ellis

The University of Melbourne, Australia

David K. Sewell

The University of Melbourne, Australia

Bradley J. Wolfgang

Defence Science and Technology Organisation,
The University of Melbourne, Australia

The relationship between attention and visual masking was investigated in a cued detection task using a factorial masking manipulation. Stimuli were either unmasked, or were masked with simultaneous (integration) masks, or delayed (interruption) masks, or integration-interruption mask pairs. The cuing effects in detection sensitivity were smallest with unmasked stimuli, intermediate with single masks, and largest with integration-interruption pairs. Large cuing effects in RT were found in all stimulus conditions. The results are inconsistent with general mechanisms of contrast gain and response gain, which do not predict interactions with interruption masks. The data were modeled using the integrated system model of visual attention of P. L. Smith and R. Ratcliff (2009), which provides an account of both RT and accuracy. The model fits suggest the action of two independent attentional mechanisms: an early selection mechanism that enhances the perceptual representation of attended, noisy stimuli, and a late selection mechanism that increases the rate of information transfer to visual short-term memory. The results are consistent with a distributed, multi-locus system of attentional control.

Keywords: attention, signal detection, spatial cuing, visual masking, decision-making, diffusion process


Introduction

One of the recurring themes in the attention literature of the last decade is of a link between visual masking and attention. By a “visual mask” we mean any stimulus that disrupts or impairs the perception of a spatially and temporally proximate target stimulus. Masking researchers, at least since Michaels and Turvey (1979), have known that the effects found in some masking paradigms depend on spatial attention. This was reinforced by the discovery of object-substitution masking, which is a form of visual masking that depends strongly on the spatial distribution of attention (Enns & Di Lollo, 1997). More recently, attention researchers have shown that the magnitude of the attentional effects in some tasks depends on the use of visual masks. The attention-masking link has manifested itself in a variety of ways (Francis, 2003; Giesbrecht & Di Lollo, 1998; Ramachandran & Cobb, 1995), but we focus here on performance in the Posner (1980) spatial cuing paradigm. In this paradigm, cues are used to attract attention covertly to a particular region of the visual field while the observer maintains central fixation.

In the visual masking literature, one of the oldest and most basic distinctions is between integration and interruption masking (Kahneman, 1968). The terminology is of comparatively recent origin, but Breitmeyer (1984) attributes the ideas to Stigler in 1926. In integration masking, the target and masking stimulus fuse to form a perceptual composite whose signal-to-noise ratio is lower than that of the target in isolation. In interruption masking, the processing of the target is terminated prematurely by the presentation of a subsequent mask. Turvey (1973) used a clerk-customer analogy to characterize interruption masking: The time a store clerk can spend serving a customer will be truncated if another customer arrives while the first customer is still being served.

More formally, integration and interruption masking have been identified with so-called Type A and Type B masking functions, respectively. A masking function characterizes the magnitude of the masking effect as a function of target-mask stimulus onset asynchrony (SOA). Type A masking functions are symmetrical around an SOA of zero. Masking is maximal when target and mask are simultaneous and decreases as the temporal separation between them increases. Forward and backward masking effects (masks leading and trailing the target, respectively)
are equal and the extent of masking depends on the relative energies of the target and mask. Type B masking functions are asymmetrical. Masking is maximal when the mask trails the target by a critical SOA, typically around 80–100 ms. The backward masking effect is much stronger than the forward masking effect and the extent of masking depends more on the temporal separation between the target and mask than on their relative energies (Turvey, 1973). The difference in the SOAs at which peak masking is found is usually taken as evidence that Type A functions reflect integration and Type B functions reflect interruption—although the neural and computational basis of these processes remains a matter for debate (e.g., Breitmeyer & Ogmen, 2000).

Masking is important in at least two contemporary theories of performance in the Posner spatial cuing paradigm. One is the integrated system model of Smith and Ratcliff (2009); the other is the perceptual template model of Dosher and Lu (2000a) and Lu and Dosher (1998). Although these theories are very different in style and in the kinds of phenomena they seek to explain, an interaction between attention and visual masking is an important feature of both of them. However, they differ in the way in which they conceptualize the role of visual masks and the way masks interact with attention. Our aim in this article is try to reconcile these two contrasting accounts.

Our investigation of the relationship between attention and masking suggests that attention acts at multiple loci within the processing system. Specifically, we find evidence for both an early and a late component of attentional selection. The early component enhances the sensory representations of noisy, attended stimuli; the late component affects the efficiency with which sensory information is made available to perceptual decision-making processes. The evidence for distinct early and late components of attention sheds new light on the early versus late selection debate that dominated the attention literature for more than two decades in the middle of the last century. Our results suggest that early and late selection need not be mutually exclusive and that processes of both kinds may operate in simple visual tasks.

### Attention-mask interactions in the Posner cuing paradigm

The integrated system model of Smith and Ratcliff (2009) was developed to account for the effects of spatial cues on response time (RT) and accuracy in detection and easy discrimination tasks—that is, in tasks in which performance is limited primarily by the contrast of the stimuli rather than by the similarity of the stimulus alternatives. The key finding leading to the development of the theory was the observation that, if stimuli are well localized perceptually, cues increase detection sensitivity (accuracy) only when stimuli are backwardly masked (Smith, 2000a). If stimuli are unmasked, detection sensitivity for cued and miscued stimuli does not differ. Smith and colleagues subsequently found this mask-dependent cuing effect in a series of studies comparing performance with masked and unmasked stimuli (Liu, Wolfgang, & Smith, 2009; Smith, Lee, Wolfgang, & Ratcliff, 2009; Smith, Ratcliff, & Wolfgang, 2004; Smith & Wolfgang, 2004, 2007; Smith, Wolfgang, & Sinclair, 2004). The integrated systems model explains this dependency on backward masking by assuming that attention affects the efficiency with which stimulus information is transferred to visual short-term memory (VSTM) and that backward masks interrupt this process before it is complete. The integrated systems model thus posits a theoretical link between cuing effects and interruption masking.

In contrast to the integrated system model, Lu and Dosher’s perceptual template model is an extended signal detection model that assumes multiple sources of noise, including external noise in the display and internal noise in the observer. Lu, Dosher, and colleagues have used this model to characterize the effects of attention on accuracy in cued discrimination tasks in which varying amounts of noise is added to the display (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000; Lu, Jeon, & Dosher, 2004; Lu, Lesmes, & Dosher, 2002). The perceptual template model distinguishes the effects of different attentional mechanisms, each with its own operating characteristic or signature. The action of these mechanisms is defined in the model by attention-dependent weighting of the signal and the various sources of noise.

One of Lu and Dosher’s most well replicated findings is that there are large cuing effects when there is external noise in the display. Lee, Itti, Koch, and Braun (1999) have reported related results. The perceptual template model attributes this to the action of an attention-dependent external noise exclusion mechanism. This mechanism, which operates only in noisy displays, allows the observer to filter out noise in the display at the target location. This results in a sharper perceptual representation of the stimulus and higher sensitivity (lower contrast thresholds) for cued stimuli. Lu and Dosher also distinguished a second mechanism, stimulus enhancement, which operates in low-noise displays. In comparison to the effects of external noise exclusion, stimulus enhancement effects are smaller in magnitude and less systematic, and have been found only with peripheral but not central cues. As discussed by Lu et al. (2002), external noise in the stimulus display can be thought of as a form of integration mask. The perceptual template model thus posits a theoretical link between attention and integration masking.

In light of Lu and Dosher’s identification of an external noise exclusion mechanism, an obvious question arises, namely, whether Smith and colleagues’ mask-dependent
cuing effect is yet another manifestation of this mechanism. To test this possibility, Smith and Wolfgang (2007) carried out a study in which observers discriminated between horizontally and vertically oriented Gabor patch stimuli presented at cued or miscued locations. The patches were masked either with noise masks or pattern masks (checkerboards) that were presented simultaneously with the target or delayed by 60–90 ms. Smith and Wolfgang (2007) argued that if the cuing effect were due to external noise exclusion, it should be largest when target and mask are simultaneous and should decrease in magnitude as target and mask are temporally separated. If, however, it were due to an interruption-masking mechanism, then it should be largest when the mask trails the target by a critical SOA. The latter is what they found. In all five of their experiments, the majority of observers showed large cuing effects in sensitivity with delayed masks. The cuing effects with simultaneous masks were smaller and were found only for a minority of observers, with both noise and pattern masks. Smith and Wolfgang argued that the large cuing effect with delayed masks was evidence for an attentional mechanism that depended on external noise exclusion, as assumed by the integrated systems model of Smith and Ratcliff (2009).

Smith and Wolfgang (2007) also contrasted the weak cuing effects they found with simultaneous masks to the results from studies using unmasked stimuli in the same paradigm (Smith, 2000a; Smith, Ratcliff, et al., 2004; Smith & Wolfgang, 2004; Smith, Wolfgang, et al., 2004). None of the studies with unmasked stimuli found a cuing effect for sensitivity for any observer. Smith and Wolfgang noted that their weak cuing effect with simultaneous masks could not have been due to a mechanism that acts via interruption masking. They therefore suggested that the effect with simultaneous masks could have been due to an external noise exclusion mechanism, as proposed by Lu and Dosher, but argued that this was not the same mechanism as the one producing the large effects with delayed masks.

Smith and Wolfgang’s (2007) study tested the proposition that the cuing effects found with backward masks could be attributed to the action of an external noise exclusion mechanism and concluded that they could not be. In this article, we test the converse proposition. Specifically, we test the conjecture by Smith and Ratcliff (2009) that the cuing effects found with both integration masks and interruption masks can be attributed to a common mechanism of attention-dependent VSTM transfer. If this is so, the interaction of attention with masks of all kinds can be attributed to a single physiological and computational locus. To investigate this question, we used a factorial masking paradigm, in which stimuli were presented in one of four different masking conditions: (a) unmasked, or masked with (b) simultaneous (integration) masks, or (c) delayed (interruption) masks, or (d) both kinds of mask together. The single-locus account predicts that integration and interruption masks will interact, both with each other and with spatial attention, to determine the quality of the stimulus representation in VSTM, and hence the speed and accuracy of the decision. It therefore predicts that the largest cuing effects will be found when an integration mask and an interruption mask are combined. It also predicts that masks of different kinds will have characteristic effects on the shapes of RT distributions. The theoretical basis for these predictions is discussed in more detail subsequently (Figure 8). The predictions for RT and accuracy provide the basis for a precise, quantitative test of the single-locus hypothesis, reported later in this article.

Methods

Apparatus

Stimuli were presented on a linearized 20 in. Sony G520 Trinitron monitor driven by a Cambridge Research Systems ViSaGe framestore controlled by a Pentium P4 computer. The ViSaGe frame rate was set to 100 Hz giving a frame duration of 10 ms. Responses were made using a CRS CT6 infra-red button box and recorded by the ViSaGe, whose on-board clock provided non-interruptible, sub-millisecond timing. Stimulus presentation and response recording were controlled by software written in C++. Observers performed the task in a dimly lit laboratory at a viewing distance of 70 cm, with their viewing position stabilized by a chin-rest.

Stimuli

Observers discriminated the orientations of horizontally and vertically oriented Gabor patches (Gaussian-vignetted sinusoidal gratings), presented on a 25° square, 30 cd/m², uniform field. The mathematical form of the Gabor patches was as given by Graham (1989, p. 53). The sinusoid had a spatial frequency of 3.5 cpd, and the Gaussian envelope had a space constant (full width at half height) of 0.463°, giving a bandwidth of 0.80 octaves. Examples of the stimuli and the display configurations are shown in Figure 1. Because contrast thresholds for orthogonal discrimination are indistinguishable from those in yes-no detection (Thomas & Gille, 1979) researchers have treated the two tasks as equivalent for the purposes of drawing inferences about attention (Cameron, Tai, & Carrasco, 2002; Lee, Koch, & Braun, 1997). Consistent with this, we have found the same pattern of mask-dependent cuing effects with backward (interruption) masks in the two tasks (Smith, 2000a; Smith, Ratcliff, et al., 2004; Smith, Wolfgang, et al., 2004). For our purposes, the orthogonal discrimination task has two advantages over the yes-no task: It is relatively unbiased,
using the procedure of Eckstein, Pham, and Shimozaki (2004), by surrounding them with four $2.5' \times 8'$ white line segments, which formed the arms of a “fiducial cross” centered on the target. The fiducial cross was presented for 60 ms at the same time as the target (Figure 1). The fiducial crosses ensured that stimuli were well localized perceptually, even when presented at low contrasts without masks (see Gould, Wolfgang, & Smith, 2007; Smith & Wolfgang, 2007, for further discussion).

On any trial, there were three possible target locations, one cued and two miscued, positioned at an angular separation of $120^\circ$ on the circumference of an imaginary $3.2^\circ$ radius circle, centered on a fixation cross. A randomly chosen angle, $\alpha (0 < \alpha \leq 360^\circ)$ determined the position of the cue on the circumference of the circle on each trial. The two possible miscued locations were at $\alpha \pm 120^\circ$. The miscued locations were equidistant from the cued location and therefore should have received equivalent processing resources. The cue consisted of four, black, right-angle markers identifying the corners of a $1.8^\circ$ square centered on one of the three possible target locations. The cues were flashed for 60 ms at a cue-target SOA of 140 ms. A weakly predictive cuing manipulation was used (Gould et al., 2007; Smith, Ratcliff, et al., 2004): Stimuli were presented at the cued location on 50% of trials and at each of the possible uncued locations on 25% of trials. Masks were presented only at the stimulus location; the other possible stimulus locations were left blank.

**Procedure**

The experiment was run using the method of constant stimuli in a three-way, Contrast (5 levels) × Cue (cued, miscued) × Mask (4 levels), design. Five levels of contrast were chosen for each of the four mask conditions individually for each observer during practice to span a range of performance from near chance to near perfect. Stimulus contrast, cue condition, and mask condition were randomized across trials within blocks in each experimental session, with an equal number of trials in each of the 40 cells of the design. We chose to use a randomized design rather than a blocked design because it implies a more highly constrained mathematical model and thus results in a more stringent test of the single-locus hypothesis, as we discuss subsequently. Each of the observers served in between 6 and 8 practice sessions, the number depending on the time required to attain a stable level of performance in each of the four mask conditions. The practice sessions were followed by 21 experimental sessions, each consisting of 480 trials, yielding a total of 10,080 trials per observer (252 trials per data point). The large number of trials per data point was required to obtain stable estimates of each of the 40 pairs of RT distributions for correct responses and errors.

Observers were instructed to try to perform the task accurately but not to spend too much time on each

minimizing criterion effects, and it yields very similar distributions of RT for the two responses. This simplified the task of fitting our mathematical models.

Four different target-mask configurations were compared. In the unmasked condition, three 10 ms Gabor frames were interleaved with three blank frames. In the integration mask condition, three target frames were interleaved with three noise frames and combined by perceptual integration (Bloch’s law), following the procedure of Lu, Dosher, and colleagues. The integration mask consisted of a Gaussian-vignetted array of $2 \times 2$ pixel squares, each of which was randomly and independently set to black or white with equal probability. The same mask was used in each of the three noise frames in a trial (static noise), but different masks were used on different trials. In the interruption mask condition, three Gabor frames interleaved with blank frames were followed at an SOA of 90 ms by 10 frames of a Gaussian-vignetted, high-contrast checkerboard. In the compound mask condition, three Gabor frames interleaved with noise frames were followed at an SOA of 90 ms by 10 frames of checkerboard. The peak contrasts of the integration and interruption mask were set at 50% and 90%, respectively. These contrasts were chosen to ensure adequate masking while yielding a sufficient range of performance to produce well-behaved psychometric functions. Stimuli were localized perceptually

Figure 1. Example stimuli. The top part of the figure shows a cued, horizontal stimulus and a miscued vertical stimulus, and the localizing fiducial crosses. The interval between the onset of the cue and the stimulus was 140 ms. The bottom panel shows the sequence of frames on a compound mask trial. Each shaded rectangle represents a 10 ms frame: light rectangles are target frames; dark rectangles are mask frames. The rectangle labeled FID shows the time course of the fiducial cross. On single-mask and unmasked trials, one or other or both sets of mask frames were omitted.

The peak contrasts of the integration and interruption mask were set at 50% and 90%, respectively. These contrasts were chosen to ensure adequate masking while yielding a sufficient range of performance to produce well-behaved psychometric functions. Stimuli were localized perceptually...
decision. They were instructed to maintain central fixation for the duration of each trial but to use the cue to direct their attention to the cued location. Because the cue-target SOA was too short to permit refixation of the display, eye movements were not monitored. (On pattern mask trials, the backward mask was presented 230 ms after cue onset and terminated 330 ms later. Although this arguably is in the range when fast saccades can be initiated, our experience in observing the display was that it is not possible to evade or reduce masking by attempting to saccade to the cue. This observation is supported by our data, which show systematic effects of the pattern mask.) Each trial began, after a 3 s intertrial interval, with presentation of the fixation cross 1 s before the cue. This served as a warning signal and an instruction to maintain fixation. The stimulus and mask were presented 140 ms after the onset of the cue. Observers indicated their decisions by pressing one of two buttons on the CT6 and were given accuracy feedback auditorily by distinctive tones. Summary accuracy feedback was also provided on the visual display at the end of each 20-trial block.

Observers

Three paid undergraduate volunteers and one of the authors (R.E.) served as observers. The volunteers were naïve to the purposes of the study. All had normal or corrected-to-normal visual acuity.

Results

Detection sensitivity

We calculated the magnitude of the cuing effect in accuracy for each observer in each of the four mask conditions by converting the proportions of correct responses to vertical and horizontal stimuli, \( P_V(C) \) and \( P_H(C) \), to \( d' \) statistics using the formula

\[
d' = \frac{z[P_V(C)] + z[P_H(C)]}{\sqrt{2}},
\]

where \( z(.) \) denotes the inverse Gaussian (z-score) transformation. The factor of \( \sqrt{2} \) in the denominator of this expression puts \( d' \) measures from a discrimination task onto the same scale as those from a yes-no task (Wickens, 2002, p. 122). We then fitted Weibull functions of the form

\[
F(c) = \alpha \left\{ 1 - \exp \left[ - \left( \frac{c}{\beta} \right) ^ \gamma \right] \right\},
\]

to the empirical \( d' \) estimates as a function of contrast, \( c \), for each observer, using the method described in Appendix A. This method yields an approximate chi-square goodness-of-fit statistic which we used to quantify model fit.

To characterize the cuing effect, we compared the fits of two Weibull models for each observer: a one-function model, in which a single Weibull function was fitted to the cued and miscued data, and a two-function model, in which the cued and miscued data were fitted with different Weibull functions. A difference in sensitivity due to cuing could be inferred if the two-function model provided a better fit than the one-function model. The results of these fits are summarized in Table 1.

Although there are some individual differences apparent in Figure 2 and Table 1, as is typically found in cuing experiments with masked stimuli, the overall pattern of results is fairly clear. For unmasked stimuli, three of the four observers showed no significant cuing effect, the exception being C.L. For noise (integration) mask stimuli, all observers showed significant cuing effects. For pattern (interruption) masked stimuli, three of the four observers showed significant cuing effects, the exception being R.E. For compound masks, all observers showed very large and highly significant cuing effects. The average cuing effect across observers is summarized in the bottom row of Table 1, which shows the average change in chi-square across observers. This effect increases progressively from left to right across the table, suggesting a progressive increase in the magnitude of the cuing effect as a function of mask condition. The plots of the group mean \( d' \) values in the bottom row of Figure 2 also show this effect.

The increase in the cuing effect as a function of mask condition is clearly evident in Figure 3, in which the cuing effect, averaged across stimulus contrasts and observers, is plotted. The figure shows a minimal or nonexistent cuing effect for unmasked stimuli, a small cuing effect for noise masks, a larger cuing effect for pattern masks, and a very large cuing effect for compound masks. The first three of these results are consistent with previous findings; the new, and striking, result is the large cuing effect with

<table>
<thead>
<tr>
<th>Observer</th>
<th>Unmasked</th>
<th>Noise</th>
<th>Pattern</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.E.</td>
<td>0.46</td>
<td>8.04*</td>
<td>1.75</td>
<td>32.18***</td>
</tr>
<tr>
<td>C.L.</td>
<td>12.81**</td>
<td>26.00***</td>
<td>35.94***</td>
<td>40.69***</td>
</tr>
<tr>
<td>J.P.</td>
<td>3.29</td>
<td>8.26**</td>
<td>79.57***</td>
<td>107.32***</td>
</tr>
<tr>
<td>S.C.</td>
<td>2.92</td>
<td>11.99**</td>
<td>21.89***</td>
<td>80.42***</td>
</tr>
<tr>
<td>Group</td>
<td>4.76</td>
<td>13.56**</td>
<td>34.79***</td>
<td>65.43***</td>
</tr>
</tbody>
</table>

Table 1. Tests of the cuing effect in sensitivity: Differences in fit between a one-function and a two-function Weibull model for unmasked, noise (integration) mask, pattern (interruption) mask, and compound (both) mask conditions. Note: * \( p < 0.05 \); ** \( p < 0.01 \), *** \( p < .001 \).
Figure 2. Detection sensitivity as a function of stimulus contrast and cue condition for cued stimuli (triangles and blue lines) and miscued stimuli (squares and red lines). The four tops rows are data for individual observers; the bottom row is the group effect, averaged across observers. The individual observer psychometric functions are fitted with Weibull functions. The error bars on the individual observer data are asymptotic standard error estimates calculated from the Gourevitch and Galanter (1967) formula; the error bars on the group data are standard errors of the between-observer means.
have used an easy discrimination task, and have presented kind of experimental tasks that were used. Whereas we unsystematic effects, can be attributed to differences in the noise in the display, whereas we have found weak and large and systematic effects with external display noise by Lu and Dosher in many experiments, and attributed by them to the action of an external noise exclusion mechanism. The fact that they experiments, and attributed by them to the action of an external cuing effects with integration masks were consistent with those found here.) The significant cuing effects found here types of masks was somewhat more pronounced than those with interruption masks. Indeed, the only surprising feature of the unmasked results in Figure 2 was the significant cuing effect for C.L., as none of the observers in any of our other studies showed a cuing effect for unmasked stimuli. This includes the studies of Gould et al. (2007) and Smith and Wolfgang (2007), both of which used fiducial crosses to localize the stimuli.

Apart from this, the results for the integration and interruption masks in Figure 2 replicate those of previous studies. Smith (2000a), Smith, Ratcliff, et al. (2004), Smith and Wolfgang (2004, 2007; Smith, Wolfgang, et al., 2004). In these studies, stimuli were localized with either luminance pedestals or fiducial crosses. Indeed, the only dissociation, it seems unlikely that the large cuing effects found with compound masks can be attributed simply to an increase in the total amount of masking: When integration and interruption masks are combined, the masking effect is increased and an enhanced cuing effect occurs. This explanation does not appeal to any qualitative differences in the way different kinds of masks interact with cues, and attributes the size of the cuing effect to the net overall masking effect.

This explanation seems to us unlikely to be correct because the size of the masking effect is not a good predictor of the size of the cuing effect. If we define the magnitude of the masking effect to be the amount by which stimulus contrast must be increased to achieve the same level of performance as is obtained with unmasked stimuli, then comparison of the abscissa values in Figure 2, along with the corresponding figures in Smith and Wolfgang (2007; Figures 2 and 4), shows that integration masks mask much more effectively than do interruption masks. Substantially higher contrasts are required to achieve any target level of performance with integration masks than with interruption masks; but the magnitude of the cuing effects found with interruption masks exceeds that found with integration masks. In other words, integration masks produce the largest masking effects, but interruption masks produce the largest cuing effects. Given this dissociation, it seems unlikely that the large cuing effects found with compound masks can be attributed simply to an increase in the total amount of masking. Instead, as we discuss subsequently, the integrated system model attributes it to the effects of masks on the efficiency of VSTM transfer.

Response time

Figure 4 shows the mean RT for each observer as a function of contrast and mask condition. The bottom row of the figure shows mean RT averaged across observers. We quantified the relationship between mean RT and contrast by fitting a two-parameter Piéron’s law (power law) function of the form

\[ MRT = ac^{-\beta}, \]  

using the method described in Appendix A. Equation 3 is simpler than the standard Piéron’s law, which includes an integration-interruption mask compounds. This result is the main focus of the modeling work reported later in the article.

The weak or nonexistent cuing effect with unmasked, perceptually well-localized stimuli has been shown repeatedly in previous studies (Gould et al., 2007; Liu et al., 2009; Smith, 2000a; Smith, Ratcliff, et al., 2004; Smith & Wolfgang, 2004, 2007; Smith, Wolfgang, et al., 2004). In these studies, stimuli were localized with either luminance pedestals or fiducial crosses. Indeed, the only surprising feature of the unmasked results in Figure 2 was the significant cuing effect for C.L., as none of the observers in any of our other studies showed a cuing effect for unmasked stimuli. This includes the studies of Gould et al. (2007) and Smith and Wolfgang (2007), both of which used fiducial crosses to localize the stimuli.

Apart from this, the results for the integration and interruption masks in Figure 2 replicate those of previous studies. Smith (2000a), Smith, Ratcliff, et al. (2004), Smith and Wolfgang (2004, 2007), and Smith, Wolfgang, et al. (2004) reported moderate-to-large cuing effects with delayed pattern masks presented at SOAs of 30 to 90 ms; but there was a small minority of observers in most studies who showed no effect. Smith and Wolfgang (2007) also found that the cuing effects with integration masks were smaller than those with interruption masks. (Indeed, the difference in the cuing effects associated with the two types of masks was somewhat more pronounced than those found here.) The significant cuing effects found here with integration masks are consistent with those found with external display noise by Lu and Dosher in many experiments, and attributed by them to the action of an external noise exclusion mechanism. The fact that they have found large and systematic effects with external noise in the display, whereas we have found weak and unsystematic effects, can be attributed to differences in the kind of experimental tasks that were used. Whereas we have used an easy discrimination task, and have presented a single stimulus in an otherwise empty display, they have typically used more complex discrimination tasks and displays containing distractor stimuli. The latter are known to produce an increase in the magnitude of cuing effects (Foley & Schwarz, 1998).

The preceding discussion implies that integration and interruption masks interact with cues in qualitatively different ways, and that an enhanced cuing effect occurs when both kinds of masks are combined. Arguably, however, there is a simpler explanation for these findings, namely, that the size of the cuing effect is just a function of the total amount of masking: When integration and interruption masks are combined, the masking effect is increased and an enhanced cuing effect occurs. This explanation does not appeal to any qualitative differences in the way different kinds of masks interact with cues, and attributes the size of the cuing effect to the net overall masking effect.

This explanation seems to us unlikely to be correct because the size of the masking effect is not a good predictor of the size of the cuing effect. If we define the magnitude of the masking effect to be the amount by which stimulus contrast must be increased to achieve the same level of performance as is obtained with unmasked stimuli, then comparison of the abscissa values in Figure 2, along with the corresponding figures in Smith and Wolfgang (2007; Figures 2 and 4), shows that integration masks mask much more effectively than do interruption masks. Substantially higher contrasts are required to achieve any target level of performance with integration masks than with interruption masks; but the magnitude of the cuing effects found with interruption masks exceeds that found with integration masks. In other words, integration masks produce the largest masking effects, but interruption masks produce the largest cuing effects. Given this dissociation, it seems unlikely that the large cuing effects found with compound masks can be attributed simply to an increase in the total amount of masking. Instead, as we discuss subsequently, the integrated system model attributes it to the effects of masks on the efficiency of VSTM transfer.
Figure 4. Mean RT as a function of stimulus contrast and cue condition for cued stimuli (triangles and blue lines) and miscued stimuli (squares and red lines). The four tops rows are data for individual observers; the bottom row is the group effect, averaged across observers. The fitted functions for individual observers are two-parameter power laws (Piéron's law). The error bars on the individual observer data are standard errors of the within-observer means; the error bars on the group data are standard errors of the between-observer means.
Table 2. Tests of the cuing effect in mean RT: Differences in fit between a one-function and two-function Piéron’s law model for unmasked, noise (integration) mask, pattern (interruption) mask, and compound (both) mask conditions. Note: **p < 0.01, ***p < .001.

<table>
<thead>
<tr>
<th>Observer</th>
<th>Unmasked</th>
<th>Noise</th>
<th>Pattern</th>
<th>Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.E.</td>
<td>12.14**</td>
<td>19.75***</td>
<td>30.40***</td>
<td>11.51***</td>
</tr>
<tr>
<td>C.L.</td>
<td>19.41***</td>
<td>29.52***</td>
<td>40.85***</td>
<td>24.78***</td>
</tr>
<tr>
<td>J.P.</td>
<td>34.87***</td>
<td>67.82***</td>
<td>57.74***</td>
<td>75.15***</td>
</tr>
<tr>
<td>S.C.</td>
<td>23.86***</td>
<td>26.44***</td>
<td>72.94***</td>
<td>102.65***</td>
</tr>
<tr>
<td>Group</td>
<td>22.57***</td>
<td>35.88***</td>
<td>50.48***</td>
<td>53.52***</td>
</tr>
</tbody>
</table>

Table 2 shows that, in contrast to the sensitivity results, there were large and systematic cuing effects for all observers in all mask conditions. This result, which parallels similar results reported by Gould et al. (2007) and Smith, Ratcliff, et al. (2004), is a striking one, because it shows that cues have different effects on sensitivity and mean RT. In the attention literature, it is often implicitly assumed that RT and accuracy are equivalent dependent variables, in the sense that they provide similar pictures of the underlying processes. The results in Figure 4 and Table 2 show that this is not the case. Whereas the sensitivity data show a mask-dependent cuing effect that varies in magnitude with mask type, the RT data show a fairly uniform cuing effect in all mask conditions. This is reinforced by Figure 5, which shows a plot of mean RT averaged over observers and stimulus contrasts. Although there are some differences between mask conditions, all of them show large cuing effects. The need to understand the dissociation in the effects of cues on accuracy and RT was one of the things that motivated the development of the integrated system model of Smith and Ratcliff (2009), discussed subsequently.

**Sensory response function**

The model assumes that the information in the stimulus is represented by a sensory response function that consists of two separable parts: an amplitude function, \( r(c) \), which depends on the contrast of the stimulus, \( c \), and a temporal response function, \( \mu(t) \), which depends on the response characteristics of the visual filter. The amplitude function is assumed to be a Naka-Rushton function of the form

\[
\frac{c^{\mu}}{c^{\mu} + c^{\text{in}}}
\]

Figure 5. Cuing effect in mean RT as a function of mask condition, averaged across observers and stimulus contrasts.
(Boynton, 2005), and the constant $\rho$ describes the nonlinearity of contrast transduction in the early visual system. (The inhibition term is often written as $c_{\text{in}} = c_{0.5}$, where $c_{0.5}$ is the so-called semisaturation constant, that is, the value of contrast at which the function attains half its maximum value of 1.0. We prefer to write it in the form in Equation 4 because it decouples the effects of divisive inhibition from the effects of nonlinearity.) Equation 4 or some variant of it have been widely used to model the psychophysics and neurobiology of visual contrast sensitivity (Foley, 1994; Heeger, 1991; Kaplan, Lee, & Shapley, 1990). The temporal response function, $r(c)$, is defined as

\[
\mu(t) = \Gamma(t; \beta_{\text{on}}, n)[1 - \Gamma(t - d; \beta_{\text{off}}, n)],
\]

where $\Gamma(t; \beta, n)$ is the output of a linear filter comprised of $n$ identical exponential stages,

\[
\Gamma(t; \beta, n) = 1 - e^{-\beta d} \sum_{j=0}^{n-1} \frac{(\beta t)^j}{j!},
\]

and $d$ is the stimulus duration. The quantities $\beta_{\text{on}}$ and $\beta_{\text{off}}$ are filter time constants that determine the onset time (rise) and offset time (fall) of the filter response. The representation of Equation 5 generalizes the usual linear system model of the visual temporal response (e.g., Sperling & Weichselgartner, 1995) to allow the rise and fall times of the filter to be different. Equation 5 has low-pass filter characteristics; its effect is to transform a brief rectangular pulsed stimulus into a smooth time-varying function of the form shown in Figure 6.\footnote{Further discussion of the properties of Equation 5 may be found in Smith and Ratcliff (2009).}

The assumptions we make about masking are the simplest ones possible. We assume that the effect of backward (interruption) masking is, in Coltheart’s (1980) terms, to limit the informational persistence of stimuli; that is, masks limit the time for which stimuli information is available to later processing stages. When stimuli are unmasked they are encoded and then subject to relatively slow iconic decay. When they are backwardly masked, the mask rapidly suppresses the sensory response. When stimuli are unmasked, we assume that $\beta_{\text{on}} > \beta_{\text{off}}$ in Equation 5; when they are backwardly masked we assume that $\beta_{\text{off}} > \beta_{\text{on}}$. The resulting sensory response functions for masked and unmasked stimuli are shown in Figure 6b. We assume that masking occurs by interruption, represented in Equation 5 as multiplicative suppression. We assume that integration masks reduce the effective contrast of the stimulus and may also change the shape of the transducer function, $r(c)$, in Equation 4. We discuss the possible effects of integration masks on sensory transduction in more detail subsequently.
Visual short term memory

In the second stage of the model, the transient information in the stimulus is encoded in a durable form in VSTM. The idea that there exists a post-iconic form of VSTM, which can survive backward masking, is a well-established one, the classical study being that of Phillips (1974). More recently, Vogel, Woodman, and Luck (2006) have studied the time course of VSTM formation in detail, using backward masking and a change detection task. They obtained estimates of the time course of VSTM encoding similar to those we have obtained here. In the integrated system model, the process of VSTM formation is modeled using a so-called shunting differential equation. The distinguishing feature of shunting equations is that the function describing the stimulus information enters into the equation multiplicatively rather than additively, as occurs in the more familiar linear-system model. This gives them nice properties when modeling VSTM processes, as we discuss below. Shunting equations have been used in models of neural computation (Grossberg, 1988; Wilson & Cowan, 1973), visual masking (Ogmen, Breitmeyer, & Melvin, 2003), and luminance discrimination (Sperling & Sondhi, 1968). Indeed, the Hodgkin-Huxley equations of neural conduction are of shunting type (Tuckwell, 1988). However, the application closest to ours is that of Busey and Loftus (1994) who, like us, use a shunting equation to model the time course of VSTM formation.

The model describes the growth of the VSTM trace, $\nu(t)$, in response to the information in the stimulus. To ensure that the VSTM trace does not saturate at long exposures, we assume the trace arises as the result of an opponent process, or excitatory-inhibitory coding scheme. In its simplest form, the VSTM growth equation can be written as

$$\frac{d\nu}{dt} = \gamma_i [r(c)\mu(t)\theta - \nu(t)] - [1 - r(c)]\mu(t)\nu(t); \quad \gamma_i \in \{\gamma_A, \gamma_U\}. \tag{7}$$

In this equation $r(c)\mu(t)$ is the excitatory coefficient and $[1 - r(c)]\mu(t)$ is the inhibitory coefficient. The equation has the property that, once the visual filter output has decayed to zero, that is, when $\mu(t) = 0$, the derivative $d\nu/dt$ also goes to zero and the trace stops changing. This property of shunting equations makes them natural models of the way in which a durable VSTM trace is computed from a transient perceptual event.

The constant $\gamma_i$ in Equation 7 is an attention gain parameter, which controls the rate of VSTM growth. Minimally, we assume that gain takes one value, $\gamma_A$, when stimuli are attended and another value, $\gamma_U$, when they are unattended, with $\gamma_A > \gamma_U$. The core assumption of the model is that attention affects the rate at which stimulus information is transferred to VSTM. When stimuli are attended, the rate of transfer is rapid; when they are unattended, transfer is slow. All of the effect of attention on performance in the model occurs via changes in the gain parameter. Smith (2000a) proposed that attention affects the rate of information transfer from early visual filters to later processing stages as an explanation of the mask-dependent cuing effect. Carrasco and McElree (2001) provided empirical support for this idea at about the same time. Smith and Wolfgang (2004) subsequently developed a quantitative model of this transfer, which is a precursor of the integrated system model. The principal difference between their model and the Smith and Ratcliff (2009) model is that the later model has an explicit VSTM stage between the visual filters and decision process. The VSTM stage was added to the model to allow it to predict the shapes of RT distributions with brief, masked stimuli. The role of the VSTM stage is to preserve the information in the stimulus in a durable form for the second or so needed to make a decision. Without such a stage, if the decision process were driven directly by the decaying outputs of early visual filters, the model would predict RT distributions to low contrast stimuli that are far more skewed than those that are found experimentally (Ratcliff & Rouder, 2000).

Equation 7 may be solved by separation of variables (Smith & Ratcliff, 2009) or by introducing an integrating factor (Smith et al., 2009) to yield an expression for the time-dependent growth of the VSTM trace:

$$\nu(t) = \theta r(c) \left\{ 1 - \exp \left[ -\gamma_i \int_0^t \mu(s)ds \right] \right\}. \tag{8}$$

Equation 8 has the following important features. First, as shown in Figure 6a, VSTM trace strength grows to an asymptote. The asymptote, which we denote $\nu(\infty)$, is proportional to $r(c)$, the Naka-Rushton transduced stimulus contrast. The constant $\theta$ determines the rate at which evidence in the VSTM trace is accumulated by the decision process. It characterizes the effective information content of the stimulus and will in general depend on the similarity of the stimulus alternatives. Second, the growth to asymptote is exponential, with a rate that depends on the attention gain parameter, $\gamma_i$. The approach to asymptote depends on the area under the temporal response function, $\mu(t)$. As shown in Figure 6b, this area depends on whether or not the stimulus is backwardly masked. When stimuli are unmasked and gain is large, the final VSTM trace strength will closely approach $\nu(\infty)$. When stimuli are masked and when gain is low, the final trace strength may be less than $\nu(\infty)$. This property is central to the model’s ability to predict the mask-dependent cuing effect as we discuss in the following section.

The full model of VSTM formation described by Smith and Ratcliff (2009) is somewhat more complex than the model of Equations 7 and 8 because it assumes that the
rate of VSTM formation is jointly dependent on attention gain and the total contrast energy of the stimulus. They introduced a contrast-dependent rate term to describe the different shapes of RT distributions that are found when stimuli are presented directly against a uniform field (Gould et al., 2007) and when they are localized by a luminance pedestal (Smith, Ratcliff, et al., 2004). We found we did not need this additional complexity to model the data we present here, so we have omitted this feature from the model. 2

Decision-making

To make a decision, successive samples of the noise-perturbed VSTM trace are accumulated over time to a response criterion. As shown in Figure 6a, there are two sources of noise in the model: between-trial noise and within-trial noise. The between-trial noise is like the noise in signal detection theory. Its effect is to produce normally distributed trial-to-trial variation in the strength of the VSTM traces representing nominally equivalent stimuli. The within-trial noise is broad-spectrum Gaussian, or white, noise. Its effect is to introduce moment-to-moment perturbations in the VSTM trace. The decision process continuously accumulates successive values of the noisy trace. This leads to a representation of the accumulating evidence as a diffusion process, denoted X(t). The growth of evidence in the decision process over time is described by a stochastic differential equation (SDE) of the form 3

\[ dX(t) = \nu(t)dt + \sigma(t)dW(t). \]  

This equation describes dX(t), the random change in decision stage activation during a small time interval of width dt. The change in activation consists of two parts: a deterministic part and a stochastic part. The deterministic part, or drift, is equal to \( \nu(t) \), the strength of the VSTM trace. The term \( \nu(t)dt \) is the mean increase in trace strength during a small interval dt. The stochastic part, \( \sigma(t)dW(t) \), describes the moment-to-moment effects of noise. The process dW(t) is the differential of W(t), the Brownian motion, or Wiener diffusion, process with standard deviation \( \sigma(t) \).

The model represents decision making as diffusion between absorbing boundaries, as shown in Figure 7. Presentation of one stimulus, a vertical stimulus, say, causes the process to drift upwards; presentation of a horizontal stimulus causes the process to drift downward. Because the sequence of noise perturbations differs on each trial, the sample paths of the process, which describe the time course of evidence accumulation, are also different on every trial. On any trial, the process accumulates evidence until it reaches one of two absorbing boundaries, an upper boundary at \( a_1 \) or a lower boundary at \( a_2 \). If the first boundary reached is \( a_1 \) a "vertical" response is made; if it is \( a_2 \) a "horizontal" response is made. The time of the first boundary crossing determines the decision component of RT. The trial-to-trial variability of the process allows the model to predict errors and distributions of RT. The model of decision-making in Figure 7 is like the diffusion model of Ratcliff (1978) except that the drift rate, \( \nu(t) \), is time-dependent, because the strength of the VSTM trace changes over time. In Ratcliff’s model, the drift rate remains constant for the duration of a trial. Mathematically, the assumption that the drift rate changes over time makes the model time-inhomogeneous, whereas Ratcliff’s model is time-homogeneous. We also assume that the moment-to-moment noise entering the decision process increases in proportion to the drift to a fixed asymptotic value. This makes the diffusion term \( \sigma(t) \) time-inhomogeneous as well. Smith and Ratcliff (2009) discuss why this assumption is a necessary and appropriate one in this setting. Methods for deriving RT and accuracy predictions for time-inhomogeneous diffusion models were described by Smith (2000b).

The mask-dependent cuing effect and a conjecture about noise

According to the integrated system model, the mask-dependent cuing effect arises from an interaction between attention gain and stimulus persistence. Gain and stimulus persistence jointly determine the final VSTM trace strength. The effects of changing gain and stimulus
Figure 8. Effects of masks on VSTM trace strength, $v(t)$, as a function of masking condition and attention gain, $\gamma_A$ and $\gamma_U$, for cued and miscued conditions. For unmasked stimuli (None), stimulus persistence is long, $\beta_{\text{off}} = 20$. The VSTM traces for cued and miscued stimuli grow to the same asymptote. With interruption masks (Pattern), stimulus persistence is short, $\beta_{\text{off}} = 200$. Asymptote trace strength for cued stimuli is greater for cued than for miscued stimuli because of the higher rate of gain for attended stimuli. With integration masks (Noise), the transfer of stimulus information is slowed by the presence of noise, which is represented by a uniform 0.5-times reduction in VSTM gain. Asymptotic trace strength is greater for cued than for miscued stimuli with long stimulus persistence because of the reduction in gain. With compound integration-interruption masks (Both), the combination of low gain and short stimulus persistence produces a large difference in asymptotic trace strength. The parameters used to generate the predictions were chosen arbitrarily to illustrate the model dynamics.

When stimuli are unmasked, stimulus persistence is long, which is represented in Equation 5 by a small value of $\beta_{\text{off}}$. The VSTM trace grows rapidly to an asymptote for cued stimuli and slowly to an asymptote for miscued stimuli, but the asymptotic trace strength for the two kinds of stimuli is the same, as shown in the upper left panel of Figure 8. Under these circumstances, the model predicts shorter RTs for cued than miscued stimuli but no differences in accuracy (sensitivity). When a backward mask interrupts the stimulus (large $\beta_{\text{off}}$), the situation is different. If the mask suppresses the stimulus before the VSTM trace formation process has run to completion, cued stimuli will have an advantage because of their higher rate of attention gain (larger value of $\gamma$). As a result, more of the VSTM trace will have formed before the mask suppresses the stimulus. Consequently, the asymptotic trace strength for cued stimuli will be greater than that for miscued stimuli, as the upper right panel of Figure 8 shows. The model predicts both shorter RTs and higher accuracy for cued, backwardly masked stimuli. Smith and Ratcliff (2009) showed that this model accurately describes the RT distributions and response accuracy for the masked and unmasked data reported by Smith, Ratcliff, et al. (2004).

The interaction between gain and stimulus persistence in the upper panels of Figure 8 occurs when the time scales of stimulus persistence and VSTM trace formation are similar. When stimulus persistence is long, trace formation runs to completion; when stimulus persistence is short, trace formation is interrupted. The bottom left panel of Figure 8 shows what happens if the trace formation process is slowed, represented by a halving of the gain for cued and miscued stimuli. The effect of reducing gain is to make trace formation slow relative to the visual persistence of unmasked stimuli. Under these circumstances, cued stimuli again have an advantage, for the same reason as they do when trace formation is fast and stimulus persistence is short. The model thus predicts a cuing effect for unmasked stimuli when overall gain is low. This is the basis for Smith and Ratcliff’s (2009) prediction about the effect of imbedding stimuli in noise.

Smith and Ratcliff (2009) argued that one of the effects of noise, or of integration masking, might be to slow the process of forming a stable VSTM representation of the stimulus—or, more particularly, of those features of the stimulus needed to do the task. They argued this based on the results of Ratcliff and Smith (2010) who studied letter discrimination in dynamic noise displays. Ratcliff and Smith found that when stimuli were imbedded in dynamic noise, changes in stimulus discriminability produced large shifts in the leading edge of the RT distribution. The leading edge (the .1 quantile) characterizes the fastest responses in the distribution. In the integrated system model, a change in the leading edge is an indicator of a change in the rate of VSTM formation. A plausible interpretation of these findings is that noise slows the rate at which a representation of the stimulus is formed in VSTM. The integrated system model therefore offers an alternative interpretation of Lu and Dosher’s external noise findings. According to this interpretation, the cuing effects that are found with both integration masks and interruption masks can be ascribed to a common locus, namely, the process of VSTM transfer. Integration masks slow the process of VSTM transfer; interruption masks terminate it before it is complete. Both lead to an increase in the size of the cuing effect.

The lower right panel of Figure 8 shows the predicted effect of an integration-interruption mask compound. The effect of the integration mask is to slow the process of VSTM transfer (a uniform reduction in $\gamma_A$ and $\gamma_U$), whereas the effect of the interruption mask is to reduce stimulus persistence (a large value of $\beta_{\text{off}}$). Combining a
reduced rate of VSTM transfer with reduced stimulus persistence leads to an enhanced cuing effect. We call this the single-locus hypothesis. This hypothesis attributes the cuing effects with both integration and interruption masks to a common process of VSTM transfer. Because the model assumes that the effect of miscuing or inattention is to slow the rate of VSTM transfer, it predicts shorter RTs to cued stimuli than to miscued stimuli, regardless of mask condition. That is, it predicts an unconditional Posner effect in RT, as shown in Figures 4 and 5. Qualitatively, both of these predictions are consistent with our experimental results. We report a detailed quantitative test of the single-locus hypothesis in following section.

Model evaluation

We follow the methods of Ratcliff and Smith (2004) and Smith and Ratcliff (2009) to evaluate our models. The data for which we wish to account are the choice probabilities (response accuracy) and the distributions of RT for correct responses and errors. We summarize the information in the RT distributions using the distribution quantiles. In our evaluation we use five quantiles: the .1, .3, .5, .7, and .9 quantiles. The .1 quantile characterizes the leading edge of the distribution (the shortest RTs), the .5 quantile describes its central tendency (the median), and the .9 quantile describes its tail (the longest RTs). We use five quantiles because they suffice to characterize the shape of the distribution while being relatively insensitive to outliers.

To evaluate our models we fitted them to the quantile-averaged data. We averaged distribution quantiles across observers for correct responses and errors in each condition to obtain group RT distributions for correct responses and errors. We also averaged the choice probabilities in each condition across observers. This yielded 40 pairs of distributions of correct responses and errors (80 distributions in all) and 40 pairs of choice probabilities. Ratcliff and colleagues have shown repeatedly that parameter estimates obtained by fitting diffusion models to quantile-averaged data agree closely with the averages of parameter estimates obtained by fitting the models to individual subject data (see Smith & Ratcliff, 2009, for references and further discussion). These results suggest that the overall picture is not distorted by fitting the models to group data. We fitted our models to the quantile-averaged data by minimizing the likelihood-ratio chi-square statistic ($G^2$),

$$G^2 = 2 \sum_{i=1}^{40} \sum_{j=1}^{12} n_i \log \left( \frac{p_{ij}}{\pi_{ij}} \right),$$

using the Matlab implementation of the Nelder-Mead Simplex algorithm (fminsearch). In this equation, $p_{ij}$ and $\pi_{ij}$ are, respectively, the predicted and observed proportions in the bins bounded by the quantiles, and “log” is the natural logarithm. The inner summation over $j$ extends over the 12 bins formed by each pair of joint distributions of correct responses and errors. (There are five quantiles per distribution resulting in six bins per distribution, or 12 bins per distribution pair.) The outer summation over $i$ extends over the five stimulus contrasts in each of the two cue conditions and four mask conditions (40 distribution pairs in all). The quantity $n_i$ is the number of experimental trials in each condition. We set this to 252, the number of trials per data point per observer. This is consistent with our interpretation of the quantile-averaged distributions as the performance of an “average observer.” Because $G^2$ computed on the joint distributions depends on the relative proportions of correct responses and errors, it characterizes goodness-of-fit to the distribution shapes and the choice probabilities simultaneously.

We investigated several dozen variants of the integrated system model to try to find a description that was both parsimonious and which fitted the data well. We discuss some of these model variants in more detail subsequently, but we focus initially on four of them. These models, which are summarized in Table 3, embody specific, substantive hypotheses about the role of attention. The models are parameterized identically, except for attention gain, $\gamma$. Table 4 lists the parameters that were estimated in fitting the four models.

Model 1 is the model for masked stimuli introduced by Smith and Ratcliff (2009). The model had a pair of attention gain parameters, $\{\gamma_A, \gamma_U\}$, one for attended and one for unattended stimuli. It also had a pair of sensory response function offset rate parameters, $\{\beta_{off,1}, \beta_{off,2}\}$, (Equation 5), one for interruption-masked stimuli and one for stimuli without interruption masks. The model predicts a mask-dependent cuing effect like that shown in the top two panels of Figure 8. Mask-dependent cuing arises from an interaction between the differences in gain for attended and unattended stimuli and the differences in persistence of masked and unmasked stimuli, as discussed previously. However the model has no mechanism for characterizing the attentional effects of noise (integration) masks, and so necessarily predicts identical cuing effects for noisy and noise-free stimuli. We use this model as a baseline model against which to compare the performance of the three other models.

Model 2 tested the single-locus hypothesis. It assumed that noise (integration) masks uniformly slow the rate at

<table>
<thead>
<tr>
<th>Model</th>
<th>$G^2$</th>
<th>$df$</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 (no noise)</td>
<td>394.5</td>
<td>421</td>
<td>556.5</td>
</tr>
<tr>
<td>Model 2 (single-locus)</td>
<td>392.8</td>
<td>420</td>
<td>563.3</td>
</tr>
<tr>
<td>Model 3 (dual-locus)</td>
<td>375.1</td>
<td>420</td>
<td>545.6</td>
</tr>
<tr>
<td>Model 4 (Unconstrained)</td>
<td>375.0</td>
<td>419</td>
<td>554.0</td>
</tr>
</tbody>
</table>

Table 3. Fits of the integrated system model.
which stimulus information is transferred to VSTM. The model assumed that interruption masks produce the same interaction between gain and stimulus persistence as assumed in Model 1. We characterized the effects of noise in this model by introducing a parameter \( \phi \), which determined the reduction in VSTM transfer rate with noisy stimuli. The attention gain in this model was parameterized as \( \{ \gamma_A, \gamma_U, \phi \gamma_A, \phi \gamma_U \} \). These four parameters describe the gain for attended noiseless, unattended noiseless, attended noisy, and unattended noisy stimuli, respectively. The lower panels of Figure 8 show an example of this parameterization with \( \gamma_A = 40, \gamma_U = 20 \), and \( \phi = 0.5 \). Very similar predictions can be obtained by assuming instead that noise reduces the sensory onset rate parameter, \( \beta_{on} \), in Equation 5. According to this assumption, noise changes the temporal response of the visual filter; specifically, it slows the rate at which the function \( \mu(t) \) in Figure 6 grows to its maximum. The reason why the two versions of the model make equivalent predictions is because of the cascaded structure of the integrated system model. Slowing the rate at which sensory information becomes available perceptually has the same effect as slowing the rate at which sensory information is transferred to VSTM. We have chosen to characterize this as a change in gain for simplicity; but some readers may prefer the sensory filter interpretation. Which interpretation is preferred is immaterial for the tests of our substantive hypotheses.

Model 3 also assumed that gain varies with noise and with attention, but in a different way. This model assumed that noise increases the magnitude of the cuing effect to a greater extent than is predicted by a uniform slowing of VSTM transfer. This is as would be predicted if an external noise exclusion mechanism were acting in addition to an interruption masking mechanism. The attention gain in Model 3 was parameterized as \( \{ \gamma_A, \gamma_{U,1}, \gamma_A, \gamma_{U,2} \} \). This model assumed that gain was the same for attended noisy and noiseless stimuli, but differed for unattended stimuli. We refer to this model as the dual-locus model, because it implies some mechanism in addition to mask- and attention-dependent changes in the rate of VSTM transfer. Model 4 relaxed the assumptions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensory Response Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset rate</td>
<td>( \beta_{on} )</td>
<td>33.8</td>
<td>45.4</td>
<td>31.1</td>
<td>31.1</td>
</tr>
<tr>
<td>Offset rate (interruption)</td>
<td>( \beta_{off,1}^* )</td>
<td>150.0</td>
<td>150.0</td>
<td>150.0</td>
<td>150.0</td>
</tr>
<tr>
<td>Offset rate (no interruption)</td>
<td>( \beta_{off,2} )</td>
<td>76.2</td>
<td>52.4</td>
<td>52.4</td>
<td>51.4</td>
</tr>
<tr>
<td>Number of filter stages*</td>
<td>( n )</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>N-R inhibition (noiseless)</td>
<td>( c_{in,1} )</td>
<td>0.073</td>
<td>0.082</td>
<td>0.081</td>
<td>0.081</td>
</tr>
<tr>
<td>N-R inhibition (noisy)</td>
<td>( c_{in,2} )</td>
<td>0.483</td>
<td>0.432</td>
<td>0.405</td>
<td>0.404</td>
</tr>
<tr>
<td>N-R exponent (noiseless)</td>
<td>( \rho_1 )</td>
<td>2.015</td>
<td>1.976</td>
<td>1.985</td>
<td>1.981</td>
</tr>
<tr>
<td>N-R exponent (noisy)</td>
<td>( \rho_2 )</td>
<td>1.595</td>
<td>1.691</td>
<td>1.663</td>
<td>1.661</td>
</tr>
<tr>
<td>Attention/VSTM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSTM asymptote</td>
<td>( \theta )</td>
<td>0.749</td>
<td>0.716</td>
<td>0.737</td>
<td>0.737</td>
</tr>
<tr>
<td>Gain (attended, noiseless)</td>
<td>( \gamma_{A,1} )</td>
<td>31.5</td>
<td>57.3</td>
<td>90.0</td>
<td>89.8</td>
</tr>
<tr>
<td>Gain (unattended, noiseless)</td>
<td>( \gamma_{U,1} )</td>
<td>21.2</td>
<td>24.5</td>
<td>39.7</td>
<td>39.5</td>
</tr>
<tr>
<td>Gain (attended, noisy)</td>
<td>( \gamma_{A,2} )</td>
<td>31.5</td>
<td>51.5</td>
<td>90.0</td>
<td>90.0</td>
</tr>
<tr>
<td>Gain (attended, noisy)</td>
<td>( \gamma_{U,2} )</td>
<td>21.2</td>
<td>22.4</td>
<td>32.1</td>
<td>32.0</td>
</tr>
<tr>
<td>Decision process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision criterion</td>
<td>( a )</td>
<td>0.112</td>
<td>0.119</td>
<td>0.120</td>
<td>0.121</td>
</tr>
<tr>
<td>Diffusion coefficient*</td>
<td>( \sigma(\tau) )</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Drift variability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmasked, cued</td>
<td>( \eta_1 )</td>
<td>0.292</td>
<td>0.269</td>
<td>0.287</td>
<td>0.289</td>
</tr>
<tr>
<td>Unmasked, misced</td>
<td>( \eta_2 )</td>
<td>0.325</td>
<td>0.296</td>
<td>0.295</td>
<td>0.297</td>
</tr>
<tr>
<td>Integration, cued</td>
<td>( \eta_3 )</td>
<td>0.254</td>
<td>0.256</td>
<td>0.262</td>
<td>0.264</td>
</tr>
<tr>
<td>Integration, misced</td>
<td>( \eta_4 )</td>
<td>0.286</td>
<td>0.302</td>
<td>0.322</td>
<td>0.325</td>
</tr>
<tr>
<td>Interruption, cued</td>
<td>( \eta_5 )</td>
<td>0.307</td>
<td>0.277</td>
<td>0.300</td>
<td>0.300</td>
</tr>
<tr>
<td>Interruption, misced</td>
<td>( \eta_6 )</td>
<td>0.459</td>
<td>0.405</td>
<td>0.408</td>
<td>0.410</td>
</tr>
<tr>
<td>Both, cued</td>
<td>( \eta_7 )</td>
<td>0.291</td>
<td>0.295</td>
<td>0.300</td>
<td>0.300</td>
</tr>
<tr>
<td>Both, misced</td>
<td>( \eta_8 )</td>
<td>0.476</td>
<td>0.336</td>
<td>0.567</td>
<td>0.567</td>
</tr>
<tr>
<td>Postdecision processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean postdecision time</td>
<td>( T_{er} )</td>
<td>0.335</td>
<td>0.356</td>
<td>0.336</td>
<td>0.336</td>
</tr>
<tr>
<td>Postdecision time range*</td>
<td>( s_\tau )</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 4. ** denotes a fixed parameter. Estimated model parameters of the integrated system model. Time dependent parameters scaled in seconds. For Model 1, \( \gamma_{A,2} = \gamma_{A,1} \), and \( \gamma_{U,2} = \gamma_{U,1} \). For Model 2, \( \gamma_{A,2} = \phi \gamma_{A,1} \) and \( \gamma_{U,2} = \phi \gamma_{U,1} \), with \( \phi = 0.898 \). For Model 3, \( \gamma_{A,2} = \gamma_{A,1} \).
of Model 3, and allowed the gain for attended noiseless, unattended noiseless, attended noisy, and unattended noisy stimuli all to be different. The set of gain parameters for this model is denoted \( \{ \gamma_{A,1}, \gamma_{U,1}, \gamma_{A,2}, \gamma_{U,2} \} \).

We evaluated the models using the Bayesian information criterion (BIC) model selection statistic. The BIC is defined as

\[ \text{BIC} = G^2 + m \log N, \]

where \( m \) is the number of parameters estimated in fitting the model and \( N \) is the number of observations used in the calculation of \( G^2 \). The BIC is a penalized likelihood statistic, which penalizes a model in proportion to its number of free parameters. Like the widely used Akaike information criterion (AIC; e.g., Smith, 1998a), the preferred model is the one with the smallest BIC. The BIC differs from the AIC in that its penalty term depends on the sample size, which gives it better properties than the AIC. Specifically, it is less prone to favor complex models with increases in sample size. The BIC values for each of the four models are shown in the right-hand column of Table 3. The degrees of freedom associated with each model are the number of degrees of freedom in the data minus the number of free parameters estimated in fitting the model. There are 440 degrees of freedom in the data (40 conditions \( \times \) 11 degrees of freedom per distribution pair); the number of free parameters was 19 for Model 1, 20 for Model 2 and 3, and 21 for Model 4.

As Table 4 shows, the parameters can be grouped into parameters of the sensory response function, parameters of the attention/VSTM selection process, parameters of the decision process, and parameters describing the post-decisional components of RT. The sensory response parameters were the onset and offset rates in Equation 5, \( \beta_{on} \) and \( \beta_{off} \), the number of cascaded stages in the filter, \( n \), and the exponent and divisive inhibition term of the Naka-Rushton function in Equation 4, \( \rho \) and \( c_{in} \). We assumed a single onset rate for the sensory response function in all conditions, but different offset rates depending on whether an interruption mask was used, as previously discussed. Because the predictions of cascade models are relatively insensitive to the number of stages in the filter (see Smith, 1995, for further discussion) we set \( n = 3 \) arbitrarily after verifying that the fits remained essentially unchanged with other choices. We allowed the Naka-Rushton exponent and divisive inhibition parameters to vary as a function of display noise (i.e., whether or not an integration mask was used). This reflects the different contrast dependencies in the psychometric functions in Figure 2. Following the work of Boynton (2005) and Lee et al. (1999), we also investigated models in which either the exponent or the divisive inhibition term or both varied as a function of attention in one or more conditions. We discuss these models later in the article.

The VSTM trace formation process is characterized by the attention gain parameters, \( \gamma \), and a parameter, \( \theta \), in Equation 8, which describes the mapping of contrast into trace strength. The decision process is characterized by two main parameters: the decision criterion, \( a \), and the diffusion coefficient, \( \sigma^2(\epsilon) \). The decision criterion determines the amount of evidence needed for a response, whereas the diffusion coefficient determines the variability of the sample paths in the accumulation process in Figure 7. We constrained the criteria for vertical and horizontal responses to be equal (\( a_1 = -a_2 \) in Figure 7) after inspection of the RT distributions and associated accuracy of the two responses. We also assumed observers use the same criteria for all stimulus types. This assumption is implied by our use of a design that randomized stimulus contrasts and mask conditions within blocks. Although it is plausible that observers use different criteria for different stimulus types in a blocked design, it is much less plausible that they change their criteria from trial to trial in response to changes in the stimulus. This was our reason for using a randomized rather than a blocked design, because it implies a more highly constrained mathematical model. One exception to this is that observers may use different criteria for attended and unattended stimuli. Indeed, some researchers have suggested that the Posner effect in detection RT may arise because observers set lower criteria at cued locations (Pashler, 1998; Sperling, 1984; Sperling & Dosher, 1986). We tested this possibility with our data and found that model fit was not improved with different criteria for cued and misused stimuli. Smith and Ratcliff (2009) found the same result for the Smith, Ratcliff, et al. (2004) data and the Gould et al. (2007) data. Of course, this does not falsify criterion setting explanations in other contexts, especially in tasks in which accuracy is high and speed of responding is stressed; but in this task it does not appear that the Posner effect in RT was due to the use of different criteria at attended and unattended locations.

In addition to these parameters, the VSTM trace formation process is characterized by a between-trials variance parameter, \( \eta \). In Ratcliff’s (1978) diffusion model, \( \eta \) describes the variance in drift between trials. This parameter allows the model to predict slow errors. In the integrated system model, \( \eta \) is identified with trial-to-trial variation in the encoded value of \( r(c) \), the Naka-Rushton transduced stimulus contrast in Equation 5. In tasks in which stimulus discriminability is low and accuracy is stressed, like ours, error RTs are typically longer than correct RTs. The model attributes this to the fact that the encoded stimulus information varies in quality randomly from trial to trial. On trials in which the encoded stimulus information is low, RTs are longer and errors are more likely. This joint dependence of RT and accuracy on the quality of the stimulus encoding results in slow errors. More discussion of this and other features of diffusion models can be found in Ratcliff and Smith (2004).

Our decision model is like the model of Ratcliff (1978) except for the fact that we assume that the drift in the...
The integrated system model has between 19 and 21 free parameters, depending on the parameterization of attention gain. Although the number of parameters is fairly large, they enter into the model in a highly constrained way, and the degree of data reduction the model affords is very substantial (approximately one parameter per 22 degrees of freedom in the data). In contrast, a purely atheoretical description, using Weibull, ex-Gaussian, or similar distributions, would require a total of 240 parameters (3 per distribution) and would give no account of choice probabilities. The integrated system model is an order of magnitude more economical in its parameterization than such descriptions.

The model fits in Table 3 show that all four models provided a reasonable account of the patterns of RT and accuracy in our empirical data. The \( G^2 \) statistics for the models are similar to one another, but there is nevertheless a clear pattern apparent in the table. In comparison with Model 1, which has no mechanism for predicting the effects of integration masking, the single-locus model (Model 2), produced essentially no improvement in fit. This is despite the fact that, qualitatively, it can predict the interaction between integration and interruption masks found empirically, as shown in Figure 8. The largest improvement in fit is found with Model 3. Like Model 2, this model had three free gain parameters, but it assumed that the magnitude of the miscuing effect was greater for integration-masked stimuli. Model 4, which allowed all four of the gain parameters to vary freely, produced no further improvement in fit over Model 3. Consistent with this, the BIC values lead to Model 3 being selected as the best of the four models. This is the model that provides the best combination of parsimony and fit. We call this model the dual-locus model because it reflects the pattern of performance that might be expected if integration masks interact with cues somewhere other than at the point of VSTM transfer, possibly at the level of the sensory response function in Figure 6.

Figure 9 shows a quantile probability plot of the fit of Model 3. The quantile probability plot is a parametric plot that shows how the shapes of the RT distributions and the choice probabilities depend on stimulus contrast and other experimental conditions. In such plots, selected quantiles of the RT distribution are plotted against the choice probabilities for correct responses and errors. Figure 9 shows the five quantiles used in fitting the data. The .1 quantile characterizes the fastest responses in the distribution (its leading edge); the .5 quantile characterizes the typical responses (the median), and the .9 quantile characterizes the slowest responses (the distribution tail). For each stimulus condition there are two RT distributions, one for correct responses (green symbols) and one for errors (purple symbols), and a choice probability (the probability of a correct response to that stimulus contrast). If the probability of a correct response at a given contrast is \( p \), the quantiles of the distribution of correct responses are plotted on the \( y \)-axis against \( p \) on the \( x \)-axis; the quantiles of the error distribution are similarly plotted against \( 1 - p \). Each pair of distributions for each level of contrast is plotted in a similar way. In the resulting plot, distributions of correct responses appear (usually) to the right of the 0.5-point on the \( x \)-axis and the distributions of errors appear on the left. The two innermost points in each panel are the distributions for the most difficult stimulus (the lowest contrast) and the two outermost points are the distributions for the easiest stimulus (the highest contrast).

The effects of the experimental manipulations on RT and accuracy are represented compactly in plots of this kind. The effect of cues on RT is represented by a vertical shift of the set of quantiles for miscued stimuli relative to those for cued stimuli. The effect of cues on accuracy is represented by a change in the horizontal extent of the plot. Any reduction in accuracy with miscuing appears as a compression of the range of \( x \)-axis values (most pronounced in the Both, Miscued condition in the panel at the lower right of Figure 9). The relative speed of correct responses and errors is reflected in the left-right asymmetry of the plot. If RTs for correct responses and errors were the same, the plot would be symmetrical around its vertical midline. It is clear from the figure that
Figure 9. Fit of the dual-locus model (Model 3). On the y-axis are the quantiles of the RT distributions; on the x-axis are the choice probabilities for correct responses and errors. The data are shown as colored symbols; the fitted values are shown as continuous curves. The lines on the graph are, in ascending order, the predicted 0.1, 0.3, 0.5, 0.7, and 0.9 distribution quantiles. The corresponding empirical quantiles are shown as circles, squares, diamonds, inverted triangles and upright triangles. The five points to the left of the 0.5-point on the x-axis (purple symbols) in each panel are the error response distributions for the five levels of stimulus contrast; the five points on the right of the 0.5-point (green symbols) are the distributions of correct responses.
error RTs are substantially longer than RTs for correct responses, as is typically found when discriminability is low and accuracy is stressed (Luce, 1986; Ratcliff & Smith, 2004). We have reported this pattern previously in other studies using versions of this paradigm (Gould et al., 2007; Smith, Ratcliff, et al., 2004). In the model, the relative speed of correct responses and errors is controlled by the drift variability parameter, $\eta$.

Figure 9 shows that the model does a fairly good job of accounting for the RT distributions and response accuracy in each of the experimental conditions. In evaluating the quality of the fit, it should be recognized that the RT distributions for errors at high stimulus contrasts are estimated empirically with low reliability, as are tail (.9) quantiles in all conditions, especially for errors. There are few errors made to easy stimuli, so these distributions are based on only a small numbers of trials. Estimates of the .9 quantile are similarly variable because of the sparseness of observations in the distribution tail. Consequently, these features of the data should be accorded relatively little weight in evaluating the fit. The main misses in Figure 9 are that the model underestimates accuracy at the lowest level of contrast in the Both, Cued condition and, to a lesser extent, in the Noise, Miscued condition. There is also some tendency for the model to miss the leading edge (.1 quantile) of the error distributions at high contrasts in the None, Cued, the None, Miscued, and the Noise, Miscued conditions. However, this is probably an artifact of quantile averaging, arising from the fact that faster observers typically produce more errors to high-contrast stimuli, and so contribute disproportionately to the quantiles of these distributions. Apart from these misses, the model provides a reasonably good characterization of the effects of cues and masks of different kinds on RT and accuracy. Model 3, like the other models in Table 3, had an interruption-mask dependent mechanism of VSTM transfer. Without such a mechanism—obtained by setting $B_{off,1} = B_{off,2}$—the fit markedly worsened: $G^2 = 501.7$, BIC = 655.2.

### Individual observer fits

To ensure our conclusions about the relative performance of Model 2 and Model 3 were not an artifact of fitting to group data, we refitted the models to the data from individual observers. These fits are summarized in Table 5. In these fits we allowed the interruption-masked sensory response offset parameter, $B_{off}$, to vary freely, as this seemed to better capture the individual differences among observers. As expected, the $G^2$ statistics for the individual fits are substantially larger than the fits to the group data, because quantile averaging smooths some of the variability that is present in individual data (see Smith, Ratcliff, et al., 2004, for a similar comparison.) Nevertheless, the model captures the main features of the cuing and masking effects for each observer. The important result in Table 5 is that the dual-locus model shows a systematic advantage for all observers. Our conclusions from fitting group data are thus also supported at the individual observer level.

### Discussion

Our main findings were that the integrated system model of Smith and Ratcliff (2009) provided a good account of the effects of cues and masks on RT and accuracy, but that the single-locus model was not the best model of the underlying attentional processes. Given that the Cue × Mask interaction in accuracy in Figure 3 appears to be captured by an interaction in VSTM trace strength like the one in Figure 8, the comparative failure of the single-locus model initially seems puzzling. However, the reason for its failure becomes clearer from close inspection of Figure 8.

According to the single-locus account, noise (integration) masks slow the rate at which a representation of the stimulus is formed in VSTM. This interacts with differences in gain ($\gamma A$ vs. $\gamma U$) for attended and unattended stimuli to produce differences in the asymptotic VSTM trace strength for cued and miscued stimuli. When no interruption mask is used, this results in small-to-moderate differences in $\nu(t)$ for cued and miscued stimuli and a commensurate cuing effect in accuracy and RT. The RT effect includes a change in the leading edge of the RT distribution that, in the integrated system model, is an index of the rate of VSTM formation. When an interruption mask is also used, the suppression of the stimulus by the mask produces large differences in $\nu(t)$ for cued and miscued stimuli. The changes in the leading edge of the RT distribution should be fairly similar in the noise mask condition and the compound mask condition, because noise has the same effect on the rate of VSTM formation in the two conditions.

Although there is some evidence of the predicted slowing in the mean RT for noise-masked stimuli in Figure 5, the effect is not reflected in the leading edge of the RT distribution; nor, indeed, is it reflected consistently in the pattern of mean RT in Figure 5. The mean RT for the compound (both) mask condition shows the predicted...
large cuing effect, but the mean RT for the cued, compound mask condition was the same as that for the unmasked and pattern-masked conditions, rather than the same as that for the noise-masked condition, as predicted by the single-locus account. The reason why the single-locus model did not provide any improvement in fit over the baseline model (Model 1), was because it was unable to account for both the differences in asymptotic VSTM trace strength and the rate of VSTM formation with a single mechanism. In contrast, the dual-locus assumed a difference in the magnitude of the cuing effect depending on whether stimuli are masked by noise. Although the model has no explicit mechanism to explain why this difference should arise, it is consistent with Lu and Dosher’s hypothesis of an external noise exclusion mechanism (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000; Lu et al., 2004, 2002), because external noise exclusion produces an increase in the signal-to-noise ratio of stimuli entering the decision process. As the simultaneous action of an interruption-masking-dependent mechanism was required to account for our data, we refer to this model as the dual-locus model.

We are cautious about invoking the classical early-late selection dichotomy in discussing contemporary vision experiments, because the distinction is tied to a particular historical and theoretical context. Nevertheless, we believe that in the present case, the distinction is illuminating. According to early selection, attention affects the quality of the sensory information available to subsequent stimulus identification and decision stages. In Treisman’s (1964) formulation of early-selection theory, the encoded representations of unattended stimuli are “attenuated.” Although this idea was originally framed in the context of attentional filtering in dichotic listening studies, the general claim is that the output of early sensory mechanisms is influenced, or modulated, by attention. In contrast, according to late selection theory, attention has its effect only later in processing, at the point at which an encoded stimulus becomes available for subjective report. In Norman’s (1968) formulation of late selection theory, attention determines whether transient stimulus activity in the memory system passes through a selective filter into conscious awareness.

Viewed in these terms, the integrated system model of Smith and Ratcliff (2009) conceives of attention as acting somewhat analogously to a late selection mechanism. In Figure 6, the transient information in the sensory response function, \( r(c) \mu(t) \), is encoded in durable form in VSTM under the control of attention. The VSTM trace, \( v(t) \), is the basis for the observer’s decision and subjective report. Changes in the attention gain, \( \gamma \), affect the quality of the trace, \( v(t) \), but have no effect on the time course or amplitude of the sensory response, \( r(c) \mu(t) \). To the extent that the model conceives of stimulus identification as a multi-stage process, involving initial stimulus encoding followed by VSTM selection and decision-making, with the effects of attention confined to the latter stages of the process, it can be conceived of as a form of late selection theory, with a particular, well-specified computational form. The single-locus model assumes that simultaneous noise masks and backward pattern masks both interact with attention at the point of VSTM selection. In contrast, the dual-locus model assumes that noise masks have their effect at some other point in the system—by implication, at the level of the sensory response, \( r(c) \mu(t) \). The dual-locus model thus admits the possibility of an additional, early selection mechanism.

**Contrast gain and response gain**

In the light of recent research showing attentional modulation in early visual pathways (e.g., Martinez-Trujillo & Treue, 2002; McAdams & Maunsell, 1999; Reynolds, Pasternak, & Desimone, 2000; Treue & Martinez-Trujillo, 1999; Williford & Maunsell, 2006), the idea of early attentional selection may seem uncontroversial. Such attention-dependent changes early in the visual system could be taken as prima facie evidence for some kind of early selection, although their implications for physiological performance remain a matter of interpretation. An early selection mechanism can be realized computationally as either a contrast gain or a response gain mechanism (Boynton, 2005; Reynolds & Heeger, 2009). Contrast gain and response gain models seek to characterize the effects of attention on the early stages of the visual contrast response, often using a Naka-Rushton function like that assumed in the integrated system model (Equation 4). In contrast gain models, attention produces a uniform increase in contrast sensitivity. The contrast response function, or transducer function, for such models can be written as

\[
\gamma_{\text{cont}}(c) = \frac{\gamma_c^o}{\gamma_c^o + c_{in}}; \quad \gamma_i \in \{\gamma_A, \gamma_U\},
\]

(12)

where \( \gamma_i \) is the attention gain. (Because we can define \( \gamma'_i = \gamma_i^o \) it is immaterial whether we write the gain term as \( \gamma(c)^o \) or \( \gamma_i(c)^o \), that is, whether we view the gain change as applied before or after the nonlinearity.) In response gain models, attention produces a uniform increase in the visual contrast response. The transducer function for response gain models can be written as

\[
\gamma_{\text{resp}}(c) = \frac{c^o}{c^o + c_{in}}; \quad \gamma_i \in \{\gamma_A, \gamma_U\}.
\]

(13)

In contrast gain models, an increase in gain results in a compression of the function on the x-axis, whereas in response gain it results in an expansion of the function on the y-axis. In logarithmic coordinates, these effects appear as a leftward shift and an upward shift, respectively.

Contrast gain and response gain models can both be realized through shunting interactions in early visual
paths, with the same kinds of dynamics that are assumed in the VSTM trace formation process (Equation 7). In either instance, we assume that attention modulates the effects of external noise, just as is assumed in Lu and Dosher’s perceptual template model. Contrast gain is obtained if the noise modulation occurs at the same stage of processing that leads to the Naka-Rushton contrastresponse function; response gain is obtained if it applies afterwards. Observe first that the contrast response function, \( r(c) \), of Equation 4 can be obtained as the asymptotic solution of the shunting equation

\[
\frac{dr}{dt} = c^o \mu(t)[1 - r(t)] - c_{in} \mu(t) r(t).
\]  

(14)

In this equation, \( \mu(t) \) is again the temporal response function of Equation 5. The asymptotic solution of this equation is

\[
r(\infty) = \frac{c^o}{c^o + c_{in}}.
\]  

(15)

which is Equation 4. Contrast gain is obtained by assuming that external display noise, denoted \( n \), acts as an additional source of inhibition and that the noise and stimulus are differentially affected by attention:

\[
\frac{dr}{dt} = \gamma_{1} \{c^o \mu(t)[1 - r(t)] - c_{in} \mu(t) r(t)\} - n \mu(t) r(t).
\]  

(16)

The asymptotic solution of this equation is

\[
r(\infty) = \frac{c^o}{c^o + c_{in} + \frac{n}{\gamma}}.
\]  

(17)

This equation has similar properties to the Foley (1994) masking model. When there is no external noise in the display (\( n = 0 \)) the contrast response is the same for attended and unattended stimuli. When there is noise present the log contrast response function is shifted to the right by an amount that depends on the gain. Large gain attenuates the effects of noise and reduces the rightward shift, that is, contrast gain.

Response gain is obtained if the inhibitory effects of noise act after the Naka-Rushton nonlinearity and normalization. We denote the time-dependent response in this hypothetical stage as \( r'(c) \). The primed notation indicates that noise acts at a stage of processing after the one involved in the computation of \( r(c) \) in Equation 4. The rate of growth in this stage is described by the shunting equation

\[
\frac{dr'}{dt} = \gamma_{1} \{r'(c) \mu(t)[1 - r'(t)] - [1 - r'(c)] \mu(t) r'(t)\} - n \mu(t) r'(t),
\]  

(18)

whose asymptotic solution is

\[
r'(\infty) = \left( \frac{c^o}{c^o + c_{in}} \right) \left( \frac{1}{1 + \frac{n}{\gamma}} \right).
\]  

(19)

When there is no external noise in the display, the contrast response is again the same for attended and unattended stimuli. When there is noise present the contrast response is divisively reduced by the noise. Large gain attenuates the divisive reduction, that is, response gain. Both of these models differ from the contrast gain and response gain models studied elsewhere in the literature in that they predict a cuing effect in accuracy only when there is external noise in the display, in agreement with our results and those of Dosher, Lu, and colleagues.

Contrast gain and response gain have provided a framework within which to study attention-dependent changes in both firing rates in cells in visual cortex (Williford & Maunsell, 2006) and visual contrast sensitivity (Huang & Dobkins, 2005; Ling & Carrasco, 2006). Such comparisons have tended not to be decisive, with some studies finding support for contrast gain (Li & Basso, 2008; Martinez-Trujillo & Treue, 2002; Reynolds et al., 2000), some for response gain (McAdams & Maunsell, 1999; Treue & Martinez-Trujillo, 1999), and some for a combination of both (Williford & Maunsell, 2006; see Reynolds & Heeger, 2009, for a review). Moreover, there is ambiguity about how these representations should be related to psychophysical measures of contrast sensitivity, because changes in the psychometric function, defined as a change in the proportion of correct responses as a function of contrast, will not behave in the same way as does changes in \( d' \) for the same data.

A much more fundamental limitation of such models is that they have no mechanism to explain the interaction with backward masking. Any effect of attention predicted by either contrast gain or response gain should be found unconditionally; it should not vary as a function of whether the stimulus is followed by an interruption mask. Although these models can be extended in a natural way to account for the effects of integration masking as we have just shown, their lack of explicit temporal dynamics means that they cannot account for attentional effects that depend on mask timing, like those we have reported here. Consequently, some other mechanism, separate from, and in addition to, attentional modulation of the visual contrast response, is needed to account for the effects of backward masking. Potentially, these effects could be attributed to the action of some kind of attention-dependent temporal window (Smith et al., 2004) or spatiotemporal template (Lu et al., 2004). The integrated system model attributes them to attention dependencies in the efficiency of VSTM transfer.

The picture that emerges from our results is that contrast gain and response gain cannot be the whole
attention story, although they may explain some part of the interaction with integration masks. Earlier, we modeled this as a difference in the rate of VSTM transfer for noisy and noiseless miscued stimuli. However, a more direct test of contrast gain and response gain is possible. We investigated this by allowing the parameters of the Naka-Rushton function in Table 4 to vary for noisy (integration masked) miscued stimuli. For contrast gain, we allowed the divisive inhibition term, \( c_{in} \), to differ for cued and miscued noisy stimuli. For response gain, we allowed the amplitude of the Naka-Rushton contrast-response function, \( r(c) \), to differ. The resulting model fits are shown in Table 6. In these model fits, the sensory response function offset parameter, \( \beta_{off} \), again varied as a function of whether an interruption mask was used. That is, these models assume that the cuing effect is a function jointly of early attentional modulation of the visual contrast response and late attentional dependencies in the rate of VSTM transfer. Unlike the previous Model 3 (the dual-locus model), the models in Table 6 assumed only two free attention parameters, \( \gamma_A \) and \( \gamma_U \), whose values were the same for noisy and noise-free stimuli.

Table 6 shows that contrast gain and response gain models result in similarly good fits and both fit around about as well as Model 3 (Table 3, group fit and Table 5, individual fits). This is unsurprising, because the VSTM trace in the integrated system model depends on the product of stimulus quality and the rate of VSTM transfer. Consequently, a low-quality stimulus and a high rate of transfer can result in the same VSTM trace strength as a high-quality stimulus and a low rate of transfer. The model fits reported in Tables 3 and 5 assumed that attention changes the rate of VSTM transfer for noisy, miscued stimuli, but has no effect on the visual contrast response function. The fits in Table 6 assumed instead that attention changes the properties of the visual contrast response function for noisy, miscued stimuli, but has no effect on the rate of VSTM transfer. Either of these possibilities results in an interaction with integration masks like the one found in the data. In either instance, the picture that emerges from our modeling is of a dual-locus process in which attention has both an early and a late component. The early component affects the quality of the sensory representation when stimuli are masked by an integration mask or when there is external noise in the display. The late component affects the efficiency of the transfer of stimulus information to VSTM prior to its suppression by an interruption mask.

We have characterized the two components of attention in the dual-locus model as an early and a late selection mechanism. A related characterization was suggested by Prinzmetal, McCool, and Park (2005), who proposed that attention could act either via a process of channel enhancement or of channel selection. In their model, channel enhancement is an increase in the quality of stimulus information at attended locations, whereas channel selection is an increase in the efficiency with which the available stimulus information is used to make a decision. These mechanisms have similarities with the two components of attention we have identified here. However, they identified the two mechanisms with the action of voluntary and reflexive orienting mechanisms, whereas we have identified them with the interaction with visual masks of different kinds.

### Table 6

<table>
<thead>
<tr>
<th>Observer</th>
<th>Contrast Gain</th>
<th>Response Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( G^2 )</td>
<td>( G^2 )</td>
</tr>
<tr>
<td></td>
<td>BIC</td>
<td>BIC</td>
</tr>
<tr>
<td>R.E.</td>
<td>1335.2</td>
<td>1327.6</td>
</tr>
<tr>
<td></td>
<td>1514.2</td>
<td>1506.6</td>
</tr>
<tr>
<td>C.L.</td>
<td>777.4</td>
<td>777.7</td>
</tr>
<tr>
<td></td>
<td>956.4</td>
<td>956.7</td>
</tr>
<tr>
<td>J.P.</td>
<td>944.2</td>
<td>943.7</td>
</tr>
<tr>
<td></td>
<td>1123.2</td>
<td>1122.8</td>
</tr>
<tr>
<td>S.C.</td>
<td>896.1</td>
<td>869.6</td>
</tr>
<tr>
<td></td>
<td>1066.6</td>
<td>1048.6</td>
</tr>
<tr>
<td>Group</td>
<td>372.4</td>
<td>371.6</td>
</tr>
<tr>
<td></td>
<td>551.4</td>
<td>542.1</td>
</tr>
</tbody>
</table>

Table 6. The bottom row (Group) shows fits to the quantile-averaged group data. All fits with 21 free parameters (df = 419).
encoding variability is increased by interruption masking and increased by miscuing. There is no strong evidence that it changes appreciably with integration masking, but the increase with interruption masking is as previously reported by Smith, Ratcliff, et al. (2004). The increase in encoding variability with miscuing suggests the presence of a trial-to-trial source of variability in attentional effects which is not represented explicitly in the integrated system model, but whose effects appear indirectly in the estimates of $\eta$. The finding is noteworthy in the light of the related finding by Prinzmetal, Amiri, Allen, and Edwards (1998) that one of the effects of inattention is to increase perceptual variability. They reported this in the context of judgments of phenomenal appearance; but our model fits suggest the presence of a similar attentional dependency in the information used to make psychophysical decisions.

**Mechanisms of interruption masking**

Throughout this article we have assumed that simultaneous noise and delayed patterns masks act via processes of integration masking and interruption masking respectively. The theoretical justification for this assumption is the work of Turvey (1973) who identified distinct central and peripheral components of masking, which were affected differently by noise and pattern masks. He showed that peripheral masking occurred by integration whereas central masking occurred by interruption. Our use of a compound mask comprising simultaneous noise and a delayed checkerboard was designed to engage the two components of masking identified by Turvey (1973).

Our argument that the mask-dependent cuing effect involves interruption masking relies on three pieces of evidence. The first, as discussed earlier, was Smith and Wolfgang’s (2007) finding that, in a paradigm similar to the present one, simultaneous masks produced small cuing effects and delayed masks produced large effects. We replicated this finding here, although the difference between the two kinds of masks was not as pronounced as in Smith and Wolfgang’s experiments. Nevertheless, we replicated their finding that simultaneous masks produce the largest masking effect, but delayed masks produce the largest cuing effect. We interpret this dissociation as evidence that the cuing effects found with simultaneous and delayed masks act through different mechanisms.

The second piece of evidence was a study by Smith and Wolfgang (2004) who investigated the mask-dependent cuing effect with monoptic and dichoptic masks, using a delayed checkerboard as the mask in either instance. They found that dichoptic presentation increased the size of the mask-dependent cuing effect, relative to the effect found with monoptic presentation. They argued, following Turvey (1973), that interruption masking should be maximized by dichoptic presentation, when two different stimuli arrive at visual cortex at different times via different routes, having had no opportunity to fuse in primary afferent pathways. The fact that the largest mask-dependent cuing effect was also found with dichotic presentation is again evidence for the involvement of an interruption masking mechanism of some kind.

The third piece of evidence was a study by Smith et al. (2009) who investigated the mask-dependent cuing effect with radial-frequency (RF) pattern masks. An RF stimulus is a suprathreshold contrast, luminance-balanced circle, whose radius is sinusoidally modulated. The observer’s task is to detect the presence of radial modulation, that is, to judge whether the stimulus is a true circle or has been radially deformed. Habak, Wilkinson, Zakar, and Wilson (2004) showed that an RF pattern is masked by a second, surrounding RF pattern that does not touch or occlude it. Habak, Wilkinson, and Wilson (2006) subsequently showed that RF pattern masking exhibits a very pronounced Type B masking function: There is little masking when the mask precedes the target or when the target and mask are presented simultaneously, but when the mask trails the target, there is a sharp increase in masking up to a critical target-mask SOA of 80–110 ms.

Smith et al. (2009) argued that the distinctive Type B masking function found with RF stimuli makes the RF masking paradigm an ideal test-bed for investigating the mask-dependent cuing effect. If the effect is associated with interruption masking, then it should be found with masked RF stimuli only when the mask is delayed. They investigated the detection of masked RF patterns in a Posner spatial cuing paradigm, using simultaneous masks and masks presented after a 100 ms SOA delay. They found no cuing effect with simultaneous masks but a large effect with the delayed masks. This is as would be expected if the mask-dependent cuing effect were associated with an interruption-masking mechanism.

The integrated system model of Smith and Ratcliff (2009) attributes the mask-dependent cuing effect to the action of an interruption-masking mechanism, but does not specify the mechanism by which interruption occurs. The only assumption made by the model, embodied in Equation 5 and shown in Figure 6, is that backward masks limit the time during which stimulus information is available to later processing stages. The model’s core predictions follow from this assumption alone, with no further assumptions about precisely how this occurs. The model thus does not rely on any potentially falsifiable assumptions about mechanisms of backward masking.

Francis (2000) has reviewed mechanisms of backward masking proposed in the literature and classified them into subtypes. Among them is a mechanism he called “efficient masking,” which he characterized using the analogy of adding cream to coffee in order to achieve maximum cooling. Maximum cooling is achieved by adding cream after a delay rather than immediately. This exploits the fact that the natural cooling effect is greatest when the coffee is hottest, so the coffee-cream mixture will be coolest if the coffee is allowed to cool for a time before the
cream is added. In the same way, the effects of some kinds of masks may be greatest if the mask is imposed after the visual representation of the stimulus has had an opportunity to decay rather than simultaneously. Efficient masking may thus be conceived of as a form of “interruption through integration.” That is, the underlying mechanism may be integration, but the masking effect may be increased if the mask is presented after the stimulus has had time to decay. We see efficient masking as providing a plausible account of the masking effects obtained with the stimuli we have used here and, indeed, we have used efficient masking as an alternative front end to the integrated system model and achieved very similar results. However, as noted above, the mechanism by which a backward masks interrupt stimulus processing is not a critical feature of the integrated system model.

Conclusion

In this article we investigated and sought to reconcile two alternative accounts of the relationship between visual masking and attention: the integrated system model of Smith and Ratcliff, and the perceptual template model of Dosher and Lu. The integrated system model posits a link between attention and interruption masking: Attention affects the efficiency of VSTM transfer and interruption masks control the time during which VSTM transfer can take place. The perceptual template model posits a link between attention and integration masking: Masks act as a source of external noise in the display, which is suppressed by attention. The starting point for the research reported here was a conjecture by Smith and Ratcliff (2009) that the relationship between attention and both kinds of masks could be explained by a common mechanism of VSTM transfer. Specifically, they conjectured that integration masks, or external display noise, slows the rate at which encoded stimulus information becomes available to latter processing stages. This mechanism can be represented in the integrated system model as a noise-dependent reduction in VSTM gain. We referred to this as the single-locus hypothesis because it proposes that the relationship between attention and the effects of both kinds of mask can be represented via a single gain parameter that controls the rate of VSTM trace formation.

We tested this hypothesis using a factorial masking manipulation in which stimuli were masked with integration masks, interruption masks, or both kinds of mask. Although the single-locus hypothesis provides an account of performance in the factorial masking paradigm that is qualitatively consistent with the experimental data, it did not provide the best fit to the joint accuracy and RT distribution data. Rather, the data were better described by a model which assumed a second, early source of attentional modulation for noise-perturbed stimuli, in addition to its effects on VSTM transfer. We referred to this as the dual-locus model. There are a variety of different ways in which this second mechanism can be realized computationally, which cannot be distinguished using our data. Within a shunting dynamics framework, this mechanism may act either as a noise-dependent contrast gain or response gain mechanism. Such mechanisms are distinct from the contrast gain and response gain models that have been investigated elsewhere in the single-cell and psychophysical literatures, as these mechanisms predict an unconditional enhancement of the stimulus representation by attention. The mechanisms we have discussed here lead instead to a selectively enhanced VSTM representation for attended stimuli when stimuli are imbedded in external noise. These mechanisms are qualitatively similar to the external noise exclusion mechanism in the perceptual template model.

The idea that attention may act at multiple loci—at the point at which stimulus information is encoded perceptually and at the point at which the encoded information is made available to decision-making mechanisms—is arguably lacking in parsimony. However, it is consistent with the failure to obtain a decisive resolution of the early-late selection debate in more than 50 years of modern attentional research. It is also consistent with the known physiology of the visual system, which has the appearance of a densely reticulated, quasi-modularized, network of neural structures, which is richly endowed with both feedforward and feedback connections. Within this network, a number of areas have been identified in which neural activity is susceptible to attentional modulation, both early and late in the implied series of neural stages or pathways (LaBerge, 1995, p. 100). Viewed in this light, the idea of a distributed, multi-locus system of attentional control is a physiologically plausible one.

Appendix A

Model-fitting procedures

To test for significant cuing effects in sensitivity as a function of contrast, $c$, we fitted Weibull functions, $F(c)$ of the form of Equation 2 to the $d'$ values for each observer by minimizing the approximate chi-square statistic

$$
\chi^2 = \sum_{i=1}^{10} \frac{[d'(c) - F_i(c)]^2}{\text{var}[d'(c)]},
$$

(A1)

using the Matlab routine fminsearch. In this expression, $d'(c)$ is the measured sensitivity in condition $i$ and $F_i(c)$ is the predicted sensitivity from Equation 2. The index of summation, $i$, runs over the five levels of stimulus contrast.
and the two types of attentional cue. The quantity in the denominator,
\[
\text{var}(d') = \frac{P_H(C)[1 - P_H(C)]}{2n_H \phi^2 \{z[P_H(C)]\}} + \frac{P_V(C)[1 - P_V(C)]}{2n_V \phi^2 \{z[P_V(C)]\}}, \tag{A2}
\]
is the asymptotic variance estimate of Gourevitch and Galanter (1967). In this expression, \(n_H\) and \(n_V\) are, respectively, the number of horizontal and vertical stimuli in condition \(i\) (set here to 126), and \(z(.)\) is the standard normal density function evaluated at the specified abscissa. The other quantities are as defined in the text. The factor of 2 in the denominator of Equation A2 is a reflection of \(\sqrt{2}\) in the denominator of the expression for \(d'\) in Equation 1. To test for the presence of cuing effects we compared the fit of an unrestricted model \(\{a_A, \beta_A, \gamma_A, a_U, \beta_U, \gamma_U\}\), in which separate Weibull functions were fitted to the psychometric functions for attended (cued) and unattended (miscued) conditions to the fit of a restricted model, \(\{a, \beta, \gamma\}\), in which the same Weibull function was fitted to the two psychometric functions simultaneously.

The fits of Piéron’s law were performed in a similar way, by minimizing an approximate chi-square statistic, using the square of the estimated standard error of the mean as a variance term in the denominator.

## Acknowledgments

The research in this article was supported by Australian Research Council Discovery Grants DP0558761 and DP0880080 to Philip Smith. The data were collected by Rachel Ellis for an unpublished B.Sc. (Honors) thesis at the University of Melbourne.

Commercial relationships: none.

Corresponding author: Philip L. Smith.

Email: philipls@unimelb.edu.au.

Address: Psychological Sciences, The University of Melbourne, Victoria, 3010, Australia.

## Footnotes

1The important feature of the function \(\mu(t)\) in Equation 5 is less its low-pass filtering properties than the fact that it characterizes the way stimulus information persists after stimulus offset. Other assumptions are consistent with the cascaded process structure of the model, but do not seem to be needed in the present setting. For example, Smith (1995, 1998b) showed how to combine band-pass visual filters with diffusion process decision models as a way to model transient perceptual channels.  

2Smith and Ratcliff (2009) compared two attentional models: a gain model and an orienting model. The gain model, like the model described here, assumed that attention affects the rate at which stimulus information is transferred to VSTM. The orienting model assumed that an attention window, or gate, opens to admit stimuli to VSTM. When stimuli are unattended the opening of the attention window is delayed. The orienting model is an elaboration of an earlier model described by Smith, Ratcliff, et al. (2004), which was inspired by the model of Reeves and Sperling (1986). Smith and Ratcliff also considered two different decision models: the single diffusion process model described here and a model described by Ratcliff and Smith (2004), in which the evidence for each response is accumulated as a separate total, each of which is an independent diffusion process. In this latter model, decision-making is conceived of as a race between two diffusing evidence totals, with the response depending on which of the processes first reaches a criterion. The model is like the leaky, competing accumulator model of Usher and McClelland (2001), but is simpler to work with, because it does not assume competitive interaction between accumulators. Smith and Ratcliff showed that, in the context of the integrated system model, the gain and orienting attention models and the single- and dual-diffusion decision models differed in only minor ways. For this reason, we have chosen to use the attention gain model and the single-diffusion decision model to analyze our data. Had we used any of the other models to analyze our data, our conclusions would not have altered.

3Stochastic differential equations are usually written in the differential form of Equation 9 rather than in the more familiar form involving derivatives. The highly irregular sample paths of diffusion processes means they do not possess derivatives in the usual sense, so quantities like \(dX/dt\) are not defined. Smith (2000b) discusses stochastic differential equations in the context of models of perceptual decision-making.

4When stimuli are unmasked, the integrated system model predicts that cues affect RT, but not accuracy. The RT effect occurs because the mean VSTM trace grows more slowly for miscued than for cued stimuli. Because the diffusive noise in Equation 9 grows in proportion to mean VSTM trace strength, the slower growth of the VSTM trace with miscued stimuli does not produce an increased likelihood of the process crossing the wrong boundary and making an error while the VSTM trace is still forming. Had we assumed abrupt-onset diffusive noise (i.e., \(\sigma(t) = 0.1\) for all \(t \geq 0\)), the model would predict both longer RTs and reduced accuracy with unmasked stimuli.

5The asymptotic solution is obtained by setting \(\mu(t) = k\) (constant), \(dv/dt = 0\), and solving for \(v(t)\). This is the solution obtained when stimulus persistence is sufficient to allow the VSTM trace formation to reach its asymptotic value, \(v(\infty)\).
References


Response times: Their role in Luce, R. D. (1986). 


Luce, R. D. (1986). Response times: Their role in inferring elementary mental organization. New York, NY: Oxford University Press.


