A STATISTICAL APPROACH TO VISUAL MASKING AND SPATIAL ATTENTION

A Dissertation

Presented to

The Faculty of the Department of Electrical and Computer Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

In Electrical and Computer Engineering

by

Sevda Agaoglu

December 2015

A STATISTICAL APPROACH TO VISUAL MASKING AND SPATIAL ATTENTION

Sevda Agaoglu

Approved:

Chair of the Committee,
Haluk Ogmen, Professor,
Electrical and Computer Engineering

Committee Members:

Bruno G. Breitmeyer, Professor,
Psychology

Ben H. Jansen, Professor,
Electrical and Computer Engineering

Bhavin R. Sheth, Associate Professor,
Electrical and Computer Engineering

George Zouridakis, Professor,
Engineering Technology

Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Badri Roysam, Professor, Chair,
Electrical and Computer Engineering

Acknowledgements

I would like to express my sincere gratitude to Dr. Haluk Ogmen for his support, guidance, continuous encouragement, constructive criticism, and cheering me up when I need it. He served as a perfect role model over the years, he always provided me with what I need to accomplish my goals and when I need it. I have always felt extremely fortunate to work with such an outstanding mentor.

I also extend my gratitude to Dr. Vallabh Das, who was an excellent mentor. I am grateful for his time, patience, constant support, and funding. I really appreciate his kindness and tolerance especially during my pregnancy. I am lucky to have such a considerate advisor.

I would like to thank the distinguished members of my doctoral committee, Dr. Breitmeyer, Dr. Jansen, Dr. Sheth, and Dr. Zouridakis for their invaluable contributions, feedbacks and encouragements.

I am also thankful to have such great lab mates and friends, who cherished my life in Houston with many wonderful memories. I especially thank my friends who participated in my extremely long experiments but never complained.

I also would like to thank my parents and parents-in-law for their support.

My special thanks to my husband and son, Mehmet and Sadi, who filled my life with joy, eased the problems I faced, gave me the energy to continue. I cannot thank enough to my best friend and colleague, Mehmet, who was always there with his unconditional love and support.

To my precious ones who cherished my life with unconditional love…

A STATISTICAL APPROACH TO VISUAL MASKING AND SPATIAL ATTENTION

An Abstract

of a

Dissertation

Presented to

The Faculty of the Department of Electrical and Computer Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

In Electrical and Computer Engineering

by

Sevda Agaoglu

December 2015

Abstract

A stimulus (mask) reduces the visibility of another stimulus (target) when they are presented in close spatio-temporal vicinity of each other, a phenomenon called visual masking. Visual masking has been extensively studied to understand the dynamics of information processing in the visual system. Visual spatial attention is also known to modulate information processing and transfer within the visual system. Since both processes control the transfer of information from sensory memory to visual short-term memory (VSTM), a natural question is whether these processes interact or operate independently. Here, we modeled visual masking by using a statistical framework, and used this theoretical framework along with psychophysical experiments to determine whether and how masking and attention interact.

In a psychophysical experiment, observers were asked to report the orientation of a target bar under three different masking paradigms. The distribution of response errors was modeled by using statistical mixture-models. Our results show that in all three types of masking, the reduction of a target's signal-to-noise ratio (SNR) was the primary process whereby masking occurred. We interpret these findings as the mask reducing the target's SNR (i) by suppressing or interrupting the signal of the target in para-/meta-contrast, (ii) by increasing noise in pattern masking by noise, and (iii) a combination of the two in pattern masking by structure.

Recent evidence suggests that the studies that reported interactions between masking and attention suffered from ceiling and/or floor effects. We investigated interactions between metacontrast masking and attention by using an experimental design

in which saturation effects were avoided. In these experiments, attention was controlled either by set-size or by spatial pre-cues. We examined attention-masking interactions based on two types of dependent-variables: (i) the mean absolute response errors and (ii) the distribution of signed response errors. Our results show that both the voluntary (endogenous) and reflexive (exogenous) mechanisms of attention affect observers' performance without interacting with masking. Statistical modeling of response errors suggests that attention and metacontrast masking exert their effects mainly through independent modulations of the guessing component of the mixture model. Taken together, our results suggest that visual masking and attention operate independently.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Visual Masking

Vision is an active process. We are surrounded by changing and moving stimuli and we are constantly processing information coming from an environment that includes both static and dynamic objects. One of the experimental tools used to study the dynamics of vision is visual masking. Visual masking is defined as the reduction in the visibility of one stimulus (target) by another stimulus (mask) when the mask is presented in the spatio-temporal vicinity of the target (Bachman, 1994; Breitmeyer & Ogmen, 2006). Visual masking has largely been investigated as a phenomenon reflecting the spatiotemporal dynamics of the visual system, and various models have been developed to explain its mechanistic bases (see reviews: Bachman, 1994; Breitmeyer & Ogmen, 2000, 2006; Enns & Di Lollo, 2000; Francis, 2000).

Visual masking has also been used to control the duration for which stimulus information remains available to the observer. After its offset, the stimulus registers first in a relatively large-capacity memory, known as sensory (or iconic) memory (Averbach & Sperling, 1961; Haber, 1983; Sperling, 1960). The contents of the iconic memory decay rapidly (< 1000ms). Iconic memory has been shown to have two components: *visible persistence* and *informational persistence (Coltheart, 1980)*. The visual system tends to keep a persisting residual image of a briefly flashed stimulus for a while (approximately 120 ms under daylight conditions) (Breitmeyer & Ogmen, 2006; Coltheart, 1980; Haber & Standing, 1970), and this phenomenon is known as *visible persistence*. The second component of iconic memory pertains to the information related

to stimulus (such as its shape, color, location, etc.) and it is not visible (Coltheart, 1980). Due to the limitations imposed by visible persistence, we would expect moving objects to be subject to a high amount of smear. For instance, if we take a camera shot of flowing traffic with a shutter speed that corresponds to the duration of visible persistence, the camera image will be highly blurred. However, under normal viewing conditions, we perceive a sharp and clear scene despite the smear that would be caused by visible persistence of visual stimuli. Our sharp and clear perception of a scene is the result of suppression of unwanted activity by the visual system that would otherwise create smear. The reduction of the blur of moving stimuli is known as *motion deblurring* (Burr, 1980). *Visual masking* has been proposed as the mechanism for *motion deblurring* (Chen, Bedell, & Ogmen, 1995; Ogmen, 1993; Purushothaman, Ogmen, Chen, & Bedell, 1998).

The distribution of receptors in the retina is not uniform. In order to enjoy the high resolution of the fovea, we often make saccadic eye movements (3-4 times per second) to reposition the fovea with the location of interest in the environment. These eye movements cause shifts and distortions in the entire retinal image, yet we do not perceive (or are not aware of) the changes that occur during these eye movements (e.g., Bridgeman, Hendry, & Stark, 1975). Visual masking has been proposed as the mechanism for rendering invisible the changes across saccadic eye movements and smear due to saccades (Bridgeman et al., 1975; Burr, 2004; Raymond, 1910).

1.2. Visual Memory

A subset of the contents of iconic memory is transferred to a more durable store, visual short-term memory (VSTM). That VSTM is a different memory store than iconic memory has been supported by the findings that a visual mask can interfere with the

contents of iconic memory but not with those of VSTM (e.g., Averbach & Coriell, 1961; Gegenfurtner & Sperling, 1993; Haber, 1983; Loftus, Duncan, & Gehrig, 1992; Schill & Zetzsche, 1995). Therefore, visual masking allows us to suppress the iconic image of a stimulus so that VSTM can be investigated in isolation. Given this property, visual masking has also played a significant role in the studies of visual memory.

The traditional view of VSTM is that, while it can store information for much longer times than iconic memory (a few seconds vs. a few hundred milliseconds), its capacity is severely limited. Most studies suggested a capacity limit of four to five items for VSTM (Cowan, 2000, 2005, 2010; Fukuda, Awh, & Vogel, 2010; Pasternak & Greenlee, 2005). Recent studies addressed whether VSTM stores its items in a fixed number of slots of equal resolution or uses a sharable resource that can be distributed among many items. Evidence for fixed slots came from studies of Zhang and Luck (2008) who used a statistical mixture model to decompose the distribution of errors into two components, a Gaussian distribution and a uniform distribution:

$$PDF(\varepsilon) = w_G * G(\mu, \sigma) + (1 - w_G) * U, \quad (1\text{-}1)$$

where *PDF* is the probability density function of errors, ε, in observers' responses; $G(\mu, \sigma)$ is a Gaussian distribution with mean μ and standard deviation σ; and U is a uniform distribution over the interval defining the errors. The Gaussian term represents reports of items in VSTM and the uniform distribution represents guesses. The parameter w_G models the proportion of responses from memory while $(1 - w_G)$ represents the proportion of guesses. The mean of the Gaussian represents the accuracy with which items are stored in VSTM and the inverse of the variance represents the precision with which items are stored. If VSTM is composed of a fixed number of

discrete slots and the number of items to report is increased, the proportion of guesses should remain close to zero until all the slots are filled (i.e., the capacity of VSTM is reached) and increase thereafter. If the slots are of fixed precision, then the standard deviation should remain independent of the number of items. While initial studies gave support for discrete fixed-precision representations in VSTM (Fukuda et al., 2010; Zhang & Luck, 2008), more recent studies provided data favoring the shared-resource approach (e.g., Bays, Catalao, & Husain, 2009; van den Berg, Shin, Chou, George, & Ma, 2012). Notwithstanding these differences, we note here the usefulness of this statistical modeling approach, which allows the separation of quantitative (w: proportion of items stored in memory) and qualitative ($1/\sigma$: precision with which items are stored) aspects of information encoding and storage.

1.3. Spatial Attention

Visual system is flooded with an enormous amount of information under normal viewing conditions. Only a subset of this information can be selected for further processing. Attentional mechanisms are responsible for enhancing the processing of the selected information (items, objects, etc.) and suppressing (or filtering out) the rest by allocating processing resources accordingly. The selection and filtering functions of visual attention have been investigated extensively and are well-documented (e.g., Chen et al., 2008; Gazzaley & Nobre, 2012; Palmer, 1990; Polk, Drake, Jonides, Smith, & Smith, 2008). In short, attention modulates the information transfer from sensory memory to VSTM and it has a significant role in maintenance of information in VSTM (Ogmen, Ekiz, Huynh, Bedell, & Tripathy, 2013; Reynolds & Chelazzi, 2004; Sreenivasan & Jha, 2007; Tombu et al., 2011). Moreover, two types of attentional orienting have been

identified (Cheal & Lyon, 1991; Egeth & Yantis, 1997; Jonides, 1981; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Posner, 1980; Weichselgartner & Sperling, 1987). *Exogenous* attention has often been described as controlled by the stimulus and, thereby, a reflexive mechanism. When we hear a loud bang or see a flash of light on a dark road, the visual systems automatically orients and deploys most of its resources for processing this information and acting upon it. Exogenous attention hence has a significant role in survival. *Endogenous* attention, however, is an *internal* allocation of resources to a predetermined region in the space, or to a particular feature. Due to this observer controlled nature of endogenous attention, it is slower. In other words, exogenous and endogenous attention differ from each other also in their temporal dynamics (Egeth & Yantis, 1997; Jonides, 1981; Wright & Ward, 2008).

1.4. Common-Onset Masking and "Object Substitution"

Although there have been many studies that directly or indirectly investigated the effect of attention on different types of visual masking (Argyropoulos, Gellatly, Pilling, & Carter, 2013; Boyer & Ro, 2007; Di Lollo, Enns, & Rensink, 2000; Filmer, Mattingley, & Dux, 2014; Hirose & Osaka, 2010; Ramachandran & Cobb, 1995; Shelley-Tremblay & Mack, 1999; Smith, Ellis, Sewell, & Wolfgang, 2010; Smith, Ratcliff, & Wolfgang, 2004; Smith & Wolfgang, 2007; Tata, 2002), the findings reported in these studies have been contradictory. In common-onset masking, where the target and mask stimuli are presented at the same time but the mask duration is longer (i.e., turned off after the target is gone), it has been claimed that attention plays a crucial role (Di Lollo et al., 2000; Enns & Di Lollo, 1997; Hirose & Osaka, 2010). According to this approach, the masking effect occurs due to *object substitution* via reentrant processing in the brain.

This process is claimed to be triggered by a mismatch between the reentrant visual representation of the target and mask pair and the incoming lower level activity due to only the mask (since it is presented alone after target's offset). As an experimental support for this hypothesis, Di Lollo, Enns, and Rensink (2000) showed that there is a significant interaction between mask-duration and set-size (the number of potential targets). Here, the assumption is that as set-size increases, attentional resources will have to be spread to more locations, thereby increasing the attentional load and, hence, the time it takes for attention to arrive at the target's location. On the other hand, when set-size is small, or the target just "pops out", attention quickly focuses on this location, hampering the masking effect. Following a similar logic, Tata (2002) showed strong interactions between masking strength and set-size in metacontrast masking, where the target and mask are spatially non-overlapping, can onset at different times, and have the same durations. This finding was intriguing because it implied that metacontrast and common-onset masking might have a common mechanism as opposed to what Di Lollo and colleagues claimed. However, recent evidence shows that, in common-onset masking, masking strength and set size (i.e., attention), do not actually interact, and that previous studies suffered from ceiling and/or floor effects which led to inappropriate conclusions (Argyropoulos et al., 2013). This finding has been recently replicated by using an eight-alternative forced choice task (Filmer et al., 2014), providing further evidence against object-substitution theory. Pilling et al. (2014) employed a spatial cue to directly control spatial attention, and also reported an absence of interaction between attention and common-onset masking.

Although there are many computational models of attention, its relation with masking is addressed by only a few models. We now briefly review two prominent models of attention which also addresses the relationship between visual masking and attention.

1.5. The Perceptual Template Model of Attention

Dosher and Lu developed a theoretical and experimental framework to investigate potential mechanisms of attention (Lu & Dosher, 1998). According to this framework, three distinct mechanisms of attention can be differentiated experimentally by adding varying levels of noise to the visual stimuli. The Perceptual Template Model (PTM) consists of four stages and incorporates both additive and multiplicative noise sources. The first stage is a "perceptual template", modeled as a filter tuned to the signal. This stage filters out some of the external noise that accompanies the desired signal. In the second stage, the output of the first stage is rectified and fed into a multiplicative Gaussian noise source with zero mean and a standard deviation proportional to the signal strength (i.e., its total energy). In the third stage, an independent Gaussian noise with zero mean and a constant standard deviation is added. The last stage is a standard signal detection (i.e., decision) process that is appropriate to the task and the stimuli.

PTM can differentiate three distinct attention mechanisms each of which leads to a signature behavioral improvement in perceptual tasks. These mechanisms are (i) stimulus enhancement, (ii) external noise exclusion, and (iii) multiplicative noise reduction. There are both physiological and behavioral evidence in support of these mechanisms. For instance, at the neurophysiological level, attention has been shown to increase cellular response sensitivity (Reynolds & Chelazzi, 2004; Reynolds, Pasternak,

& Desimone, 2000), to sharpen tuning curves of orientation and spatial frequency selective cells (Haenny, Maunsell, & Schiller, 1988), and to shrink neuronal receptive fields thereby excluding unwanted information through intra- or inter-layer interactions (Desimone & Duncan, 1995). At the behavioral level, attention has been associated with reduction in decision uncertainty (Palmer, Ames, & Lindsey, 1993), enhancement of the attended stimuli (Lu & Dosher, 1998; Lu, Liu, & Dosher, 2000; Posner, Nissen, & Ogden, 1978), exclusion of external noise or distractors (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 2000; Lu, Lesmes, & Dosher, 2002; Shiu & Pashler, 1994), and modulation of contrast-gain (Lee, Itti, Koch, & Braun, 1999).

Figure 1-1 illustrates the model and how these three different mechanisms of attention can be distinguished from each other based on the equivalent-input-noise (Dosher & Lu, 2000b; Lu & Dosher, 1998, 2005) method. In this method, external noise with varying magnitudes is injected to the system along with a signal. The goal is to determine the internal noise of the system by analyzing performance as a function of external noise. When performance is plotted as a function of external noise magnitude, two asymptotic regimes emerge: When external noise is much smaller than internal noise, only internal noise will limit performance. Hence, performance will be independent of external noise and it will follow a horizontal asymptote. On the other hand, when external noise is much larger than internal noise, performance will depend uniquely on external noise and will follow an oblique asymptote, which is proportional to external noise. The external noise at which performance transitions from the horizontal to the oblique asymptote provides an estimate of the internal noise of the system. By using this technique, one can distinguish between these three attentional mechanisms as follows.

Consider first the stimulus enhancement mechanism, which is obtained by amplifying the output of the perceptual template stage. This enhancement amplifies the signal and some of the external noise, since the perceptual template reduces but does not completely eliminate external noise. On the other hand, internal noise is left intact. Hence, performance along the oblique regime will not change because, here internal noise has no effect and external noise dominates, and attention amplifies not only the signal but also the external noise leaving the signal-to-noise ratio unaltered. On the other hand, one predicts an improvement for the performance in the horizontal regime: Here, although external noise is amplified, it is still negligible with respect to internal noise and the enhancement of the signal will improve the signal-to-noise ratio since the dominant noise (internal noise) remains constant whereas the signal becomes enhanced.

Another potential mechanism of attention is external noise exclusion: Attention operates on the specificity of the perceptual template (i.e., the bandwidth of the filter) so that less external noise can enter the system. In contrast to the previous case, here one predicts no change in the horizontal asymptote regime and improved performance in the oblique asymptote regime: For the horizontal asymptote regime, external noise is already negligible with respect to the internal noise. Hence, reducing external noise further does not result in an improvement in performance. On the other hand, for the oblique regime, external noise is dominant, and hence a reduction of effective external noise by attention is predicted to improve performance. In fact, another simple way to explain this effect is to highlight that reducing external noise is equivalent to shifting the performance curve as a function of external noise to the right.

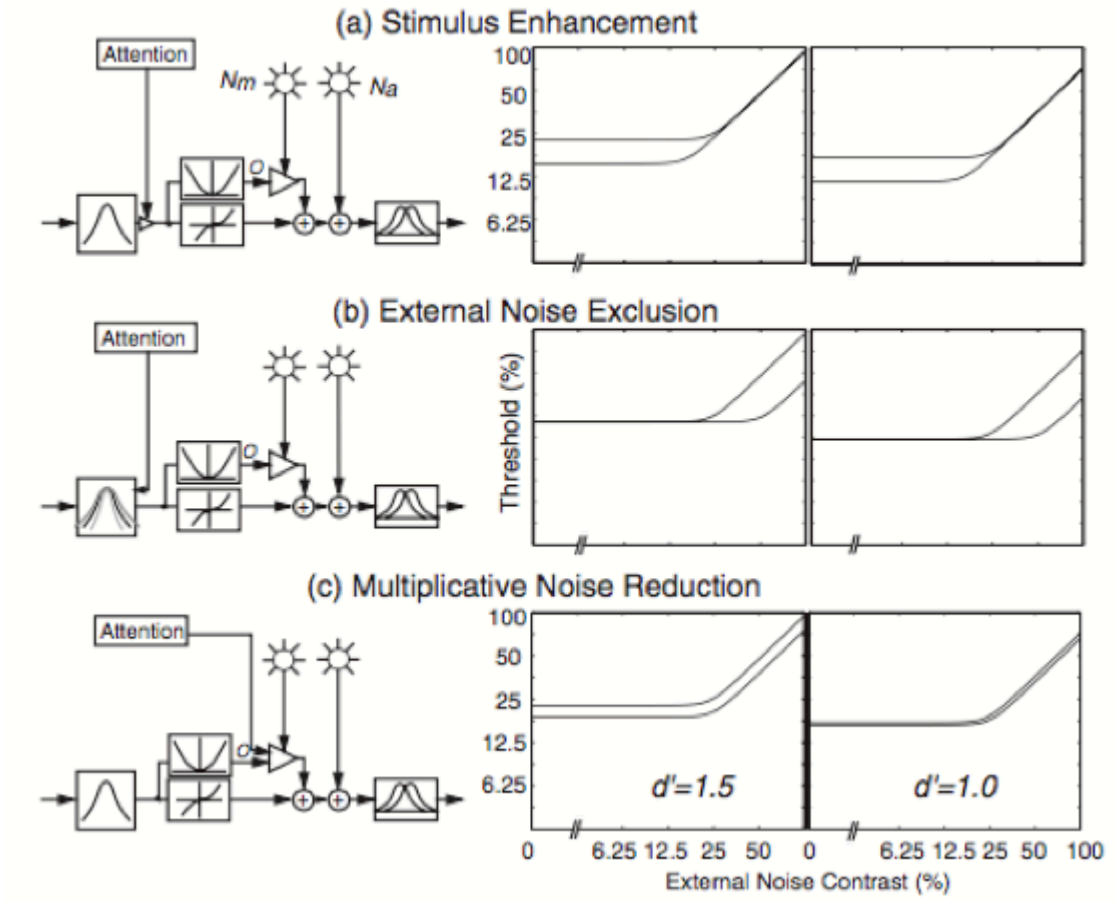

Figure 1-1 The Perceptual Template Model (Lu & Dosher, 2005). (a) Attention might enhance the signal, or (b) modulate the specificity of the perceptual template or (c) modulate multiplicative noise. See text for details. Reproduced from Lu & Dosher, 2005.

The third potential mechanism of attention that can be distinguished by this paradigm is multiplicative noise reduction. It predicts increased performance (or decreased thresholds) in both low and high external noise conditions since attention reduces the gain of the multiplicative noise generator stage. If signal enhancement and external noise reduction mechanisms operate in parallel, the resultant change in thresholds will not be distinguishable from that due to multiplicative noise reduction mechanism alone. However, this potential pitfall can be overcome by measuring

thresholds at multiple performance criterion levels (Lu & Dosher, 2004). A higher level of threshold performance will not change the effect of attention in signal enhancement or external noise exclusion mechanism but it will lead to a smaller effect of attention in multiplicative noise reduction mechanism. Therefore, Dosher and Lu suggested and employed multiple performance criteria when measuring thresholds in order to fully distinguish these three mechanisms. Although there is empirical evidence for the first two mechanisms, there have been no findings in favor of the third mechanism (Dosher & Lu, 2000a, 2000b; Eckstein, Thomas, Palmer, & Shimozaki, 2000; Lu & Dosher, 1998; Lu et al., 2002, 2000; Palmer, Verghese, & Pavel, 2000; Pestilli & Carrasco, 2005; Yeshurun & Carrasco, 1998).

As we have mentioned before, there are two broad categories of spatial cueing, namely central and peripheral cueing. Central cues are generally presented at the locus of fixation and signals the location of the target stimulus in a way that requires interpretation. For example, when an arrow is used, the observer has to interpret the direction of the arrow to infer the cued location. Central cueing is claimed to activate voluntary or endogenous attention mechanisms. Peripheral cues are generally presented at the periphery, at or close to the spatial location of the stimulus and hence they indicate the location of the stimulus directly in spatial representations without necessitating interpretive processes. These cues are said to activate the reflexive, or exogenous, attention mechanisms. Lu and Dosher (2000) found that endogenous attention works by external noise exclusion whereas exogenous attention invokes both external noise exclusion and signal enhancement mechanisms. Ling and Carrasco (2006), however,

showed that both types of attention increase contrast sensitivity in both high- and low-noise conditions.

PTM also addresses the mask-dependent cueing effects, where spatial cueing of a location increases performance only when the target stimulus is masked (e.g., (Lu & Dosher, 1998, 2000; Lu et al., 2002; Smith & Wolfgang, 2004, 2007). According to this model, masking increases the external noise. In their studies, Dosher, Lu, and colleagues used only noise masks with short target-mask stimulus onset asynchronies (SOA) (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000; Lu, Jeon, & Dosher, 2004; Lu et al., 2002). At small SOAs, target and mask falls within the temporal- integration window and, their activities fuse together (integration masking). This, as assumed by Dosher, Lu, and colleagues, effectively decreases the target SNR by increasing external noise. Other types of masks and long SOAs are not addressed by PTM. However, by making several assumptions, we can still draw predictions for metacontrast masking with a range of SOAs.

As we mentioned before, PTM predicts that external noise exclusion is the mechanism underlying endogenous attention effects. Under the external noise exclusion scenario, PTM predicts large attentional effects when external noise is large. If the mask's effect is to add noise to the stimulus, then more noise should be added when masking is strong. Accordingly, the effect of attention should be strong when masking is strong and weak when masking is weak, hence there should be interactions between attention and masking.

1.6. The Integrated System Model of Attention

Smith and colleagues developed the integrated system model (ISM) to explain spatial cueing effects on accuracy and reaction time in detection tasks (Smith & Wolfgang, 2004 – early version, no explicit VSTM layer; Smith & Ratcliff, 2009 – VSTM stage is added; Smith et al., 2010 – final version). The main assumption of the model is that attention affects the rate of information transfer from sensory level to VSTM (Carrasco & McElree, 2001). The model assumes two separate mechanisms for integration (short SOAs) and interruption (intermediate and long SOAs) masks regardless of the spatial layout of them. For integration masking, the model assumes that mask reduces the effective contrast of the visual stimulus. For interruption masking, it is assumed that the processing of the visual stimulus is terminated prematurely by the mask (i.e., informational persistence Coltheart, 1980 is interrupted) and hence, the information about the stimulus is not available for further processing stages.

The model consists of three parts: (i) Sensory response stage, (ii) VSTM stage, and (i) a sequential diffusion process for decision making (Figure 1-2). In the sensory response layer, the transient characteristics of a visual stimulus are encoded by multiple spatiotemporal filters. The sensory response function is space-time separable. The amplitude is a function of stimulus contrast. An integration mask reduces the effective contrast of the visual stimulus. The critical parameters of the temporal response filter are the rise and decay time constants of the filter depending on whether there is a delayed mask or not. When the stimulus is not masked, the activity-related stimulus is subject to relatively slower iconic decay. When the stimulus is backward-masked, the decay is more rapid.

In the VSTM layer, sensory information is accumulated into a more durable form, which is immune to visual masking (Phillips, 1974). VSTM trace formation is governed by shunting differential equations, which were previously used to model VSTM trace growth (Busey & Loftus, 1994) and other visual phenomena (Grossberg, 1988; Ogmen, Breitmeyer, & Melvin, 2003; Sperling & Sondhi, 1968; H. R. Wilson & Cowan, 1973). The rate of the VSTM trace growth is determined by an attention gain parameter. The model assumes that the attention gain parameter is larger when the stimulus is attended whereas it is smaller when the stimulus is unattended (Smith, 2000). The asymptote of the VSTM trace, which corresponds to the strength of the representation of the stimulus and determines the accuracy of detection, depends on the stimulus strength (e.g., contrast) and whether it is masked or not. Therefore, ISM also addresses the relationship between visual masking and attention. When attention gain is high and stimuli are unmasked, then the VSTM trace will quickly approach its asymptote (i.e.,$v(\infty) = \lim_{t\to\infty} v(t)$) whereas when attention gain is low and stimuli are masked, it will take more time for VSTM to reach its plateau, which will be lower than v(∞). The critical assumption here is that masking and attention strongly interact, again based on mask-dependent cueing effects found in earlier studies.

The last layer is the decision layer that is a sequential diffusion process. Noisy successive samples of VSTM trace are accumulated to make a decision. There are two different noise sources to achieve variability within and between trials. Within trial noise is a Gaussian white noise and introduces moment-to moment perturbations to the VSTM trace. Between-trial noise contributes to trial-to-trial variability. As soon as the VSTM

trace reaches one of the boundaries of the diffusion process, a decision is made and the first boundary cross determines the reaction time (RT) of a certain trial.

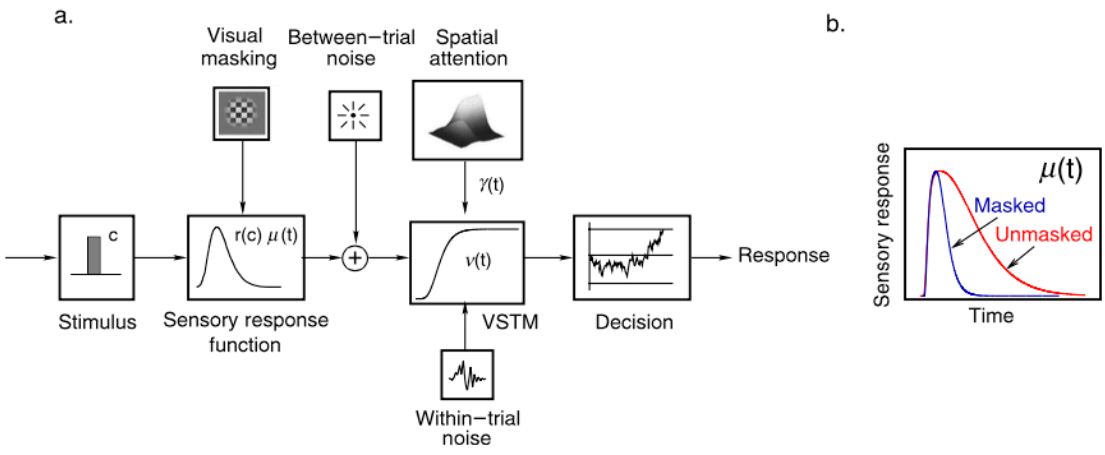

Figure 1-2 a. The Integrated System Model as in Smith et al. (2010). b. Sensory response of a stimulus. If it is masked by an integration mask, effective contrast of the stimulus is reduced. If it is masked by an interruption mask, stimulus response is truncated. Reproduced from Smith et al., 2010.

According to the integrated system model, the mask dependent cuing effect arises from an interaction between attention gain and informational persistence. When there is no mask, both the contrast and duration of the sensory response is intact and so is the VSTM strength (i.e., the asymptote). In addition, attending the stimulus (i.e., a larger attention gain) increases the rate of the VSTM trace growth but does not affect the VSTM strength. Therefore, when there is no mask, perceptual accuracy is not affected although RT is smaller for attended stimulus. When there is an integration mask, the effective contrast of the stimulus is reduced and hence, the VSTM growth is slower and it reaches a lower plateau. Integration masking and attention interact as follows: The attention gain parameter is unaffected by the presence of an integration mask when the stimulus is attended, whereas it is further reduced by integration masking when the stimulus is

unattended. This leads to a mask-dependent cueing effect in accuracy of detection. The model implicitly assumes that interruption masks do not directly interact with attention. When there is an interruption mask, the sensory response of the visual stimulus is prematurely truncated and hence, VSTM strength is low. Moreover, the presence/absence of an interruption mask does not change attention gain. When both integration and interruption masks are applied to a stimulus, suppressive effects of masks combine and mask-dependent cueing effects in both accuracy and RT become even larger.

To summarize, according to ISM, mask-dependent cuing effect found both with integration and interruption masking types can be attributed to a common process: VSTM transfer. The integration mask slows down the rate of VSTM transfer by modulating attention gain and effective contrast of the stimulus, while an interruption mask terminates it prematurely by truncating sensory response to the stimulus (Smith et al., 2010). Crucially, ISM employs interacting masking and attention mechanisms, therefore it predicts larger attentional benefits when a stimulus is masked compared to when it is unmasked. Likewise, the stronger the masking is, the larger the attentional effects will be.

1.7. General Objectives

In summary, the relations between visual masking and sensory (iconic) memory, and the relations between sensory memory and attention have been studied extensively in various contexts. The third leg of the trivet, upon which a complete understanding of visual perception in a dynamic environment can be established, is the relationship between visual masking and attention. Although there are few studies investigating this relationship for common-onset masking, the true nature of masking-attention link still

remains poorly understood due to contradicting reports. In addition, whether and how other types of masking and attention interact is still not known.

Given the importance of visual masking, both as an experimental tool and as a mechanism of normal vision, the **broad goal** of this dissertation is to further our understanding of masking phenomena and its relation to visual spatial attention. Specifically, we adopted the statistical models used to unravel the relationship between visual memory and attention, and we investigated how masking occurs, and whether and how different types of spatial attention (endogenous and exogenous) interact with visual masking.

In the first part of the study, we investigated how mask-related activities might influence the target-related ones from a statistical point of view. In contrast to existing mechanistic models (Bridgeman, 1971, 1978; Francis, 2000; Ogmen, 1993; Weisstein, 1968, 1972), our statistical approach provides a descriptive understanding of masking. We modeled the distribution of response errors of human observers in three different visual masking experiments, namely para-/meta-contrast masking, pattern masking by noise, and pattern masking by structure. Following scenarios may occur during visual masking: (i) Mask activity may "interfere" with the encoding of a target and cause decreased precision in observer's reports. (ii) Mask activity may reduce a target's signal-to-noise ratio (SNR) without interfering with its encoding precision. (iii) Decreased performance due to masking may result from the confusion or "misbinding" of a mask's features with those of the target, when they are similar as in the case of pattern masking by structure. We adopted statistical models described before, which have been used

previously in studies of visual short-term memory, to capture response characteristics of observers under masking conditions to determine which scenario occurs.

In the second part, we investigated the effect of attention on metacontrast masking, and the interactions, if any, between them by varying the set size and Stimulus Onset Asynchrony (SOA), the time difference between the target and mask onsets. Then, we utilized a similar statistical approach described above to capture response characteristics of observers whose attention is manipulated by varying set-sizes. Similar scenarios as discussed above may take place under various attention conditions except that incorrect identity binding may occur because of the distractors instead of mask elements. An important point that should be addressed here is a possible change in the 'winning model' parameters as set size changes and whether there is an interaction between model parameters obtained for different set size conditions as a function of SOA. In other words, our aim is to reveal the dependencies, if any, between the quantity and quality of information that gets through the masks and varying attentional loads.

It has been known that exogenous and endogenous cues lead to different performance characteristics as a function of Cue Target Onset Asynchrony (CTOA), the time difference between the cue and target onsets. In the last part, we investigated the timing and dynamics of spatial attention and visual masking. To do so, we took advantage of different time courses of endogenous and exogenous orienting, and we varied both SOA and CTOA. In this part, we kept the set size the same across conditions in contrast to the second part of the study. Again, we did statistical modeling based on observer's response errors and determine if there is any interaction between model parameters for different CTOA values in exogenously and endogenously cued conditions.

1.8. Specific Aims

1.8.1. Part I – A Statistical Perspective to Visual Masking

How can target and mask interactions be characterized from the statistical point of view? We tested the following hypotheses:

Hypothesis 1: The reduction of visibility of the target is due to an "interference" on target encoding by the mask activity, resulting in decreased precision in observer's reports.

Hypothesis 2: The reduced visibility of the target results from decreased signal-to-noise ratio (SNR) caused by the mask activity resulting in increased rate of guesses in subjects' responses without interfering with the encoding precision of the target stimulus.

Hypothesis 3: Masking is caused by the incorrect identity binding of a feature of the mask to the target stimulus.

1.8.2. Part II – Interactions between Spatial Attention and Visual Masking

First, we looked at the effect of set-size on metacontrast masking strength. Second, we determined whether there is an interaction between set-size and metacontrast masking by comparing performance for different set-size conditions as a function of SOA. Here we tested the following hypothesis:

Hypothesis 4: Attention and metacontrast masking are independent processes; hence, there should be no interaction between these two.

Third, we characterized observers' responses to understand how the effect of attention and the interaction, if any, between attention and masking are reflected in the

quantity and quality of information present in the brain. The hypotheses listed in Part I were also tested here with one exception. Incorrect identity binding described in *Hypothesis 3* was now due to distractor items rather than mask elements.

1.8.3. Part III – Temporal Dynamics of the Effect of Endogenous and Exogenous Attention on Visual Masking

Attentional allocation can be controlled by changing set size (Part II). However, there are two limitations of controlling attentional allocation this way. First, it does not allow us to investigate the temporal dynamics of attentional allocation. Second, since observers have to attend to the entire display at the beginning of each trial and the target is indicated by the onset of a mask, the task employs both endogenous and exogenous attention. It has been known that endogenous and exogenous attention have different temporal dynamics and hence, one cannot tease apart their relative contribution to stimulus encoding accuracy and precision in the brain with such a task. Here, we directly controlled the allocation of attention by using a spatial cue indicating the location of the target among distractors. In separate experiments, we used endogenous and exogenous cues. This allowed us to examine the timing of attentional allocation and the difference between endogenous and exogenous attention. We investigated whether there is an interaction between cue timing and visual masking by comparing performances for different CTOA conditions as a function of SOA for exogenous and endogenous cueing. Then, we tested the following hypotheses:

Hypothesis 5: Endogenous attention and metacontrast masking are independent processes; hence, there should be no interaction between the two.

Hypothesis 6: Exogenous attention and metacontrast masking are independent processes; hence, there should be no interaction between the two.

In order to characterize the target-mask interactions, if any, we used the same statistical models described above for different cue-types and CTOA conditions. After determining the statistical model that explains best the behavioral data, we looked at whether there is an interaction between model parameters obtained for different cue types and CTOA values. Finally, comparison of the best models and associated model parameters for different cue types allowed us to determine statistical signatures of endogenous and exogenous cueing.

Chapter 2. A Statistical Perspective to Visual Masking[1]

2.1. Introduction

Visual masking is defined as the reduction in visibility of one stimulus (target) by another stimulus (mask) when the mask is presented in the spatio-temporal vicinity of the target (Bachmann, 1984; Breitmeyer & Ogmen, 2006). Visual masking has largely been investigated as a phenomenon reflecting the spatiotemporal dynamics of the visual system, and various models have been developed to explain its mechanistic bases (reviews: Bachmann, 1984; Breitmeyer & Ogmen, 2000, 2006; Enns & Di Lollo, 2000; Francis, 2000). Visual masking has also been used as an experimental tool, often to control the duration for which stimulus information remains available to the observer. After its offset, the stimulus registers first in a relatively large-capacity memory, known as iconic or sensory memory (Averbach & Sperling, 1961; Haber, 1983; Sperling, 1960). The contents of the iconic memory decay rapidly, within a few hundred milliseconds. A subset of the contents of iconic memory is transferred to a more durable store, visual short-term memory (VSTM). That VSTM is a different memory store than iconic memory has been supported by the findings that a visual mask can interfere with the contents of iconic memory but not with those of VSTM (e.g., Averbach & Coriell, 1961; Gegenfurtner & Sperling, 1993; Haber, 1983; Loftus et al., 1992; Schill & Zetzsche, 1995). Given this important criterion, visual masking has played a significant role in studies of visual memory.

[1] All of the findings reported in this chapter have been already published (Agaoglu, et al., 2015).

The traditional view of VSTM is that, while it can store information for much longer times than iconic memory (few seconds vs. few hundred milliseconds), its capacity is severely limited. Most studies suggested a capacity limit of 4 to 5 items for VSTM (Cowan, 2000, 2005, 2010; Fukuda et al., 2010; Pasternak & Greenlee, 2005). Recent studies addressed whether VSTM stores its items in a fixed number of slots of equal resolution or uses a sharable resource that can be distributed among many items. Evidence for fixed slots came from studies of Zhang and Luck (2008) who used a statistical mixture model to decompose the distribution of errors into two components, a Gaussian distribution and a uniform distribution:

$$PDF(\varepsilon) = w_G * G(\mu, \sigma) + (1 - w_G) * U, \quad (2\text{-}1)$$

where *PDF* is the probability density function of errors, ε, in observers' responses; $G(\mu, \sigma)$ is a Gaussian distribution with mean μ, and standard deviation σ; and *U* is a uniform distribution over the interval defining the errors. The Gaussian term represents reports of items in VSTM and the uniform distribution represents guesses. The parameter w_G models the proportion of responses from memory while $(1 - w_G)$ represents the proportion of guesses. The mean of the Gaussian represents the accuracy with which items are stored in VSTM and the inverse of the variance represents the precision with which items are stored. If VSTM is composed of a fixed number of discrete slots and the number of items to report is increased, the proportion of guesses should remain close to zero until all the slots are filled (i.e., the capacity of VSTM is reached) and increase thereafter. If the slots are of fixed precision, then the standard deviation should remain independent of the number of items. A second version of this model assumes that resources can be shared among the slots; in this case, the standard deviation should

remain independent of the number of items when set-size exceeds the number of slots. While initial studies gave support for discrete fixed-precision representations in VSTM (Fukuda et al., 2010; Zhang & Luck, 2008), more recent studies provided data favoring the shared-resource approach (e.g., Bays et al., 2009; van den Berg et al., 2012). Notwithstanding these differences, we note here the usefulness of this statistical modeling approach, which allows the separation of quantitative (w: proportion of items stored in memory) and qualitative ($1/\sigma$: precision with which items are stored) aspects of information encoding and storage. Since a visual mask deteriorates the contents of iconic memory and hence affects what can be transferred into VSTM, our goal in this study was to characterize how masks affect the quality and quantity of information by using a similar modeling technique.

In particular, given the parameters of the statistical model in Equation (2-1), we wanted to consider the following scenarios: 1) The mask may lead to a reduction in the weight of the Gaussian term (equivalently an increase in the weight of the Uniform term, since these two add to unity) without affecting the standard deviation or the mean of the Gaussian. This case may be interpreted as the mask reducing the signal-to-noise ratio (SNR) of the target signal without affecting the encoding quality of the target[2]. Since signal and noise are intertwined in the SNR, a priori we cannot tell whether the reduction in SNR occurs via a reduction in signal strength, via an increase in noise strength, or via a combination of both. We will call this case the "SNR effect". 2) The mask's activity may

[2] Performance depends both on the strength of the target signal, captured by the weight of the Gaussian, and the encoding quality of the target signal, captured by the mean and the standard deviation of the Gaussian. When we refer to signal-to-noise ratio, we are referring to the relative weights of the target Gaussian and the uniform distributions. While one may also consider the accuracy and precision limits (i.e. the mean and the standard deviation, respectively, of the Gaussian) to stem also from noise processes, our use of noise in this manuscript refers exclusively to that underlying the relative weights of the target Gaussian and the uniform distribution.

"interfere" with the target's encoding and cause a change in the mean and/or the standard deviation of the Gaussian term. We will call this case the "interference" effect. 3) Finally, by using an extension of the aforementioned model, we will also assess whether the decrease in performance due to masking results from the confusion or misbinding of the mask's features with those of the target, when the target and mask are similar as in the case of pattern masking by structure. Since masking is not a unitary phenomenon (reviews: Bachmann, 1984; Breitmeyer & Ogmen, 2006), we separately analyzed para-/meta-contrast masking, pattern masking by noise, and pattern masking by structure (see Figure 2-1).

2.2. General Methods

2.2.1. Participants

Five observers (three naïve observers and the authors SA and MA; two female, three male) participated in the study. The age of the participants ranged from 26 to 39 years and all participants had normal or corrected-to-normal vision. The work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Experiments followed a protocol approved by the University of Houston Committee for the Protection of Human Subjects. Each observer gave written voluntary informed consent before the experiments.

2.2.2. Apparatus

Visual stimuli were created using the ViSaGe card manufactured by Cambridge Research Systems. Stimuli were displayed on a 22-in. CRT monitor. Resolution was set

to 800x600 and the refresh rate was 100 Hz. Observers were at a distance of 1 m from the screen. In order to help observers to keep a stable gaze, a fixation cross at the center of the screen and a head/chin rest were provided. Behavioral responses were recorded via a joystick. We devoted 100 trials to each stimulus onset asynchrony (SOA) separating the onset of the target from that of the mask, in order to obtain a satisfactory number of data points for statistical tests. The entire experiment required 15 sessions, with 5 separate sessions for each masking type. The order of the sessions for different masking types was randomized in order to minimize order effects. To assess baseline target visibility, twenty trials in which only the target was presented were interleaved in every session. Practice trials were run to familiarize the observers with the task and the settings of the experiments.

2.2.3. Stimuli

The target and the mask were presented at a 6-deg horizontal eccentricity in the right half of the display while the observers fixated at the center of the screen. Background luminance was 40 cd/m^2. The target was an oriented bar, 1 deg long and 0.1 deg wide (Figure 2-1a). The luminance of the target differed for each type of masking and exact values will be specified in respective sections. The mask (Figure 2-1b-d) was either a non-overlapping ring (para-/meta-contrast) or a random dot pattern (masking by noise) or 3 bars with the same dimensions as the target but varying in orientation (masking by structure).

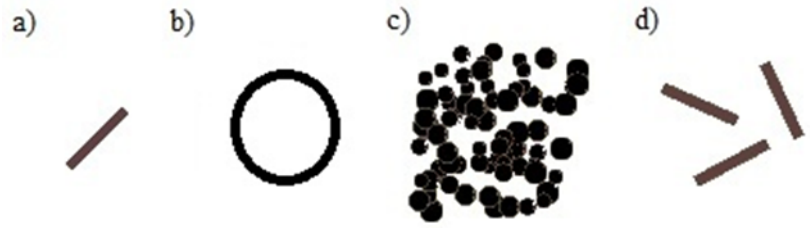

Figure 2-1 Stimuli Configurations. a) Target b) Para-/Meta-contrast Mask c) Noise Mask d) Structure Mask.

Figure 2-2 shows an example of the stimuli sequence. Each trial started with a fixation cross. The duration of the fixation cross was randomly chosen from the values between 0.5 sec and 1 sec. For positive SOA values, the target was shown first, followed by a blank interval determined by the SOA value. For negative SOA values, the order was reversed. Observers were asked to give their responses by adjusting via a joystick the orientation of a bar shown 1500 ms after the offset of the stimulus at the center of the screen. No feedback was provided to the subjects. The initial orientation of the response bar was randomly chosen among the values ranging from 0 to 179 deg. The resolution of the adjustment via joystick was 1 deg of orientation.

Para-/meta-contrast masking and pattern masking by structure typically generate Type-B masking functions (maximum masking occurs at a positive SOA value) whereas pattern masking by noise generates Type-A functions (maximum masking occurs at SOA=0) (Breitmeyer & Ogmen, 2006). Since we focused on different parts of masking functions for Type-A and Type-B, we employed different SOA values for different mask types. The durations of the target and the mask were 10 ms.

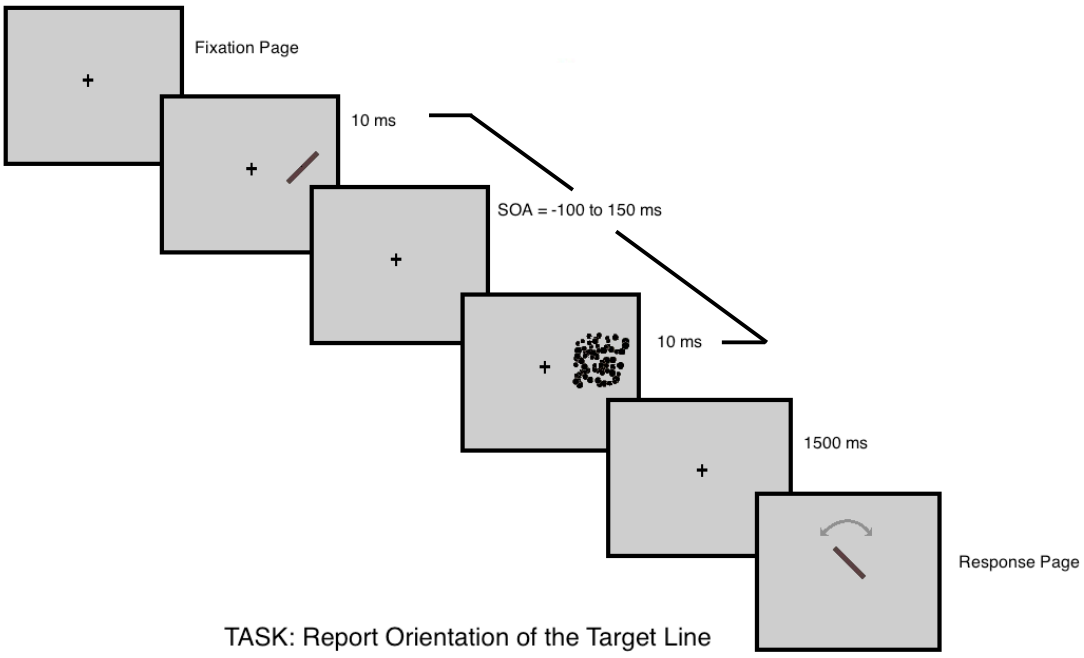

Figure 2-2 Example of sequence of stimuli. For positive SOA values (as depicted), the mask followed the target. For negative SOA values, the order was reversed. The duration of the blank interval between the target and the mask presentations was determined by the SOA.

2.2.4. Analysis

To obtain masking functions, we transformed each observer's orientation settings at each SOA for the three types of masks used here, by first computing response errors. Error values were calculated as the difference between the actual and the reported angles. Error values ranged from -90 to 90 deg. Transformed performance (Ogmen et al., 2013) was then calculated as

$$Transformed\ Performance = 1 - \frac{|Error\ Angle|}{90} \quad . \tag{2-2}$$

When the observer produces no errors, error angle will be zero, resulting in a Transformed Performance value of 1. When the observer purely guesses, the average of the absolute value of error angles will be 45 deg. The corresponding Transformed

Performance will be 0.5. Hence, transformed performance is a *linear* transform that converts errors to a probability-like measure such that transformed performance values of 0.5 and 1 correspond to chance and perfect performance, respectively.

2.2.5. Statistical Models

We adopted statistical models that have been used previously in modeling VSTM (Bays et al., 2009; Zhang & Luck, 2008). In pattern masking by noise and para-/meta-contrast masking paradigms, we analyzed two different models to explain the masking effect. The first model (Gaussian[3]) suggests that the masking effect is caused by the "interference" of the mask signal with the target signal in such a way that the encoding precision and/or accuracy for the target signal is hampered. Decreased stimulus encoding precision is reflected by the increased variability of behavioral responses. A Gaussian distribution (Figure 2-3a) is used to model this effect (the Gaussian model will be referred to as the G model). The mean of the Gaussian distribution converts to a measure of the accuracy of the system; i.e., the closer the mean to 0 the more accurate the system. The reciprocal of the standard deviation of the Gaussian distribution reflects the encoding precision of the system. In the second model, Gaussian + Uniform (the GU model), an increased guess rate caused by a drop in the target's SNR is also considered. The increased guess rate is modeled by the weight of the uniform distribution as shown in Figure 2-3b. The GU model is a weighted sum of Gaussian and Uniform distributions (Equation 2-3):

$$PDF(\varepsilon) = w_G * G(\mu, \sigma) + (1 - w_G) * U. \quad (2\text{-}3)$$

[3] One might argue that modeling a circular data with a regular Gaussian may not be appropriate. However, in early model simulations, we used both "wrapped" (or circular) Gaussians, also known as von Mises distribution, as well as regular Gaussians. We found virtually no difference between the two, and hence, we used regular Gaussians for the rest of our analyses for simplicity.

In pattern masking by structure, since the mask elements share structural properties of the target, there is a possibility to report one of the mask elements, e.g., the one that has the closest angle to the target angle or the closest location to the target location, instead of the target stimulus. In this case, the masking effect would be caused by incorrect identity binding and it is modeled by an extra Gaussian distribution shown in Figure 2-3c. If the source of this extra Gaussian component is the mask element which has the closest angle to the target angle, then the model is "Gaussian + Uniform + Closest Angle," and it will be referred to as the GUCA model. If misbinding is caused by the mask element which is closest to the target location, i.e., its nearest neighbor, then the model is "Gaussian + Uniform + Nearest Neighbor," and it will be referred to as the GUNN model. In this case, the PDF is a weighted sum of target Gaussian and non-target Gaussian distributions and the Uniform distribution (Equation 2-4) (Bays et al., 2009). We call this model "the Misbinding model," given by

$$PDF(\varepsilon) = w_T G(\mu_T, \sigma_T) + w_{NT} G(\mu_{NT}, \sigma_{NT}) + (1 - w_T - w_{NT}) U(-\pi/2, \pi/2), \quad (2\text{-}4)$$

where subscripts T and NT denote target and non-target parameters, respectively. Note that the models come from an embedded family, i.e., two or more PDFs are embedded into a family of PDFs that are indexed by one or more parameters (Kay, 2005). Given the different number of parameters in each model, an adjustment for the number of parameters is needed for comparing model performances. For instance, the Misbinding model contains the components of the GU model, and the same relationship is valid for the GU and G models, too. These relationships allowed us to identify potential contributions of different mechanisms; if adding a new component to the model enhances

model performance (how well the model explains the experimental data), then it would imply presence of a mechanism modeled by this new component.

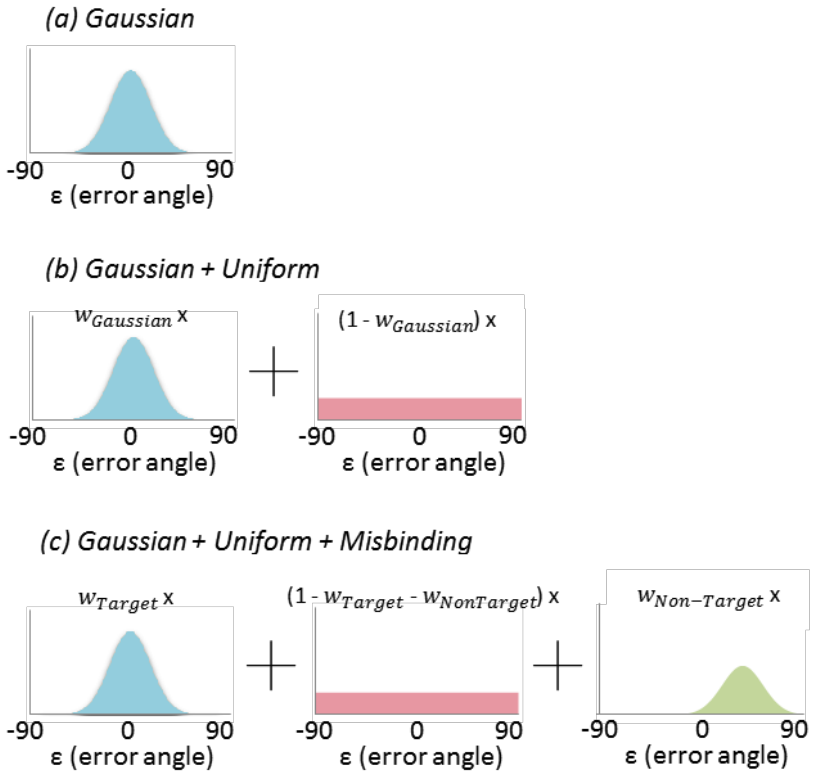

Figure 2-3 Statistical Models tested in this study. a) Gaussian b) Gaussian + Uniform c) Gaussian + Uniform + Misbinding Terms

2.2.6. Model Fitting and Model Comparison

We used two different techniques to determine model parameters and to compare different models. In the first technique, we used the Least-Mean-Squares (LMS) approach to find the best fitting parameters and the adjusted-R^2 criterion to select the model that explains the data best. The results of these analyses are presented in Appendix A. As a second approach, we used the Bayesian Model Comparison (BMC) technique (Mackay, 2004; Wasserman, 2000). We present the BMC technique and its results in the

main text of the manuscript. Overall, the two techniques produced very similar results, thereby indicating that both our parameter determination and our model selection process were robust.

Each model m_j produces a predicted error distribution p(ε|m_j,θ), where ε is a vector of observed response errors, and θ is a vector of model parameters. For each model, we calculated the likelihood of finding observed response errors, averaged over free parameters:

$$L(m_j) \triangleq p(\varepsilon|m_j) = \int p(\varepsilon|m_j,\theta)p(\theta|m_j)d\theta = \int \left(\prod_{i=1}^{N} p\left(\varepsilon_i\middle|m_j,\theta\right)\right) p\left(\theta\middle|m_j\right)d\theta, \quad (2\text{-}5)$$

where N represent the number of trials and ε_i represents the error in the i^{th} trial. It is convenient to take the logarithm of Equation (2-5) in order to compute it numerically. Equation (2-5) can be rewritten as

$$\ln L(m_j) = \ln L_{max}(m_j) + \ln[\int exp(\ln L(m_j|\theta) - \ln L_{max}(m_j))\, p\left(\theta\middle|m_j\right)d\theta], \quad (2\text{-}6)$$

where $\ln L(m_j|\theta) = \sum_{i=1}^{N} ln\ (p\left(\varepsilon_i\middle|m_j,\theta\right)$, and $L_{max}(m_j) = \max(L(m_j|\theta))$. Parameters corresponding to $L_{max}(m_j)$ can be regarded as the Maximum Likelihood Estimation (MLE) of the model parameters for model m_j. Subtracting $L_{max}(m_j)$ ensures that the exponential in the integrand is of order 1 and thereby, avoids numerical problems (Ester, Zilber, & Serences, 2015; Mackay, 2004; van den Berg, Shin, et al., 2012). Since we do not have an a priori reason to do otherwise, we used a uniform distribution over a plausible range of parameters for the parameter prior distributions (see Table 2-I). For G and GU models, the priors were a one- and two-dimensional uniform distributions, respectively:

$$p(\theta|m_j) = \prod_{t=1}^{k} \mathrm{U}(\theta_{t,min}, \theta_{t,max}), \quad (2\text{-}7)$$

where U(a,b) represents a uniform distribution over the interval [a,b], k represents the number of free parameters in the model m_j, and $\theta_{t,min}$ and $\theta_{t,max}$ represent the minimum and maximum boundaries for the t^{th} free parameter. For GUCA and GUNN models, both of which have four free parameters, the probabilities over parameter space were again uniform distributions; but the prior was not simply a hypercube with n=4, (i.e., a product of four independent uniform distributions), since not all model parameters are independent for these models. To be more specific, the sum of w_{target} and $w_{non\text{-}target}$ cannot exceed 1. If the range for w_{target} is [0,1], the corresponding range for $w_{non\text{-}target}$ can only be [0,1- w_{target}]. In other words, the support of the probability function is the triangular region consisting of the non-negative space with $w_{target} + w_{non-target} \leq 1$. Since all other parameters are independent, the joint prior for the GUCA and GUNN models can be expressed as a product of three independent uniform distributions and a triangular distribution. For the parameter μ_T, the mean of the target Gaussian, we chose $\mu_T = 0$, corresponding to a Dirac delta function as prior. This was motivated by the following: In visual masking, there is no a priori reason to expect a significant bias for the mean of the Gaussian (in contrast, in visual crowding for instance, such systematic trends can be expected (see for example: Ester et al., 2015; van den Berg, Johnson, Anton, Schepers, & Cornelissen, 2012)). Indeed, in our approach using the LMS+adjusted R^2 method, we found that the mean of the Gaussian is not significantly different from zero (see Appendix A.) Therefore, in the following analyses, the target Gaussians were centered on target orientations (i.e., zero mean in error space), which decreased the number of free parameters in all models.

Table 2-I Range of parameters used for BMC. Note that in a separate analysis for para/meta-contrast masking and pattern masking by noise, we used step sizes of 0.1 for the standard deviation of the Gaussian, and 0.002 for the weight of the Uniform but the results were not affected.

	σ_T	w_U	w_{NT}	σ_{NT}
Upper bound	50	1	1	50
Lower bound	1	0	0	1
Step size	1	0.02	0.02	1

σ_T Standard deviation of the target Gaussian
w_U Weight of the Uniform
w_{NT} Weight of the non-target Gaussian
σ_{NT} Standard deviation of the non-target Gaussian

Considering these priors, Equation (2-6) becomes

$$\ln L(m) = \ln L_{max}(m) - \sum_j^k \ln (R_j) + \ln[\int exp(\ln L(m|\theta) - \ln L_{max}(m))\, d\theta],^{4} \quad (2\text{-}8)$$

where R_j represents the size of the range for the jth free parameter. We approximated the integral by a Riemann sum with at least 50 bins in each parameter dimension (we also repeated the analysis for G and GU models with 500 bins in each parameter dimension and verified that the results with 50 bins are sufficiently robust). We refer to the performance metric given in Equation (2-8) as BMC. The difference between BMC from two different models is equivalent to the log of their likelihood ratios.

2.2.7. Analysis of Model Parameters

After determining the best model in explaining the statistics of observers' response errors, we sought to find how different model parameters change as a function of SOA. The reasoning behind this analysis was to determine which one of the scenarios listed in the Introduction section best accounts for the visual masking phenomenon. We examined the model parameters that yielded L_{max} to see how they vary as a function of

[4] For a more detailed derivation of this results, please see Appendix E.

SOA and compared the results to the masking functions in order to assess whether they correlate well with masking. We report traditional ANOVA results as well as Bayes factors for this analysis. Bayes factor analyses were done in the programming language R using the "BayesFactor" package developed by Rouder et al. (available for download at bayesfactorpcl.r-forge.r-project.org, see also for reference Rouder, Morey, Speckman, & Province, 2012). We also quantified the correlation between the model parameters and the masking strengths by calculating Pearson R coefficients. Masking strength is calculated as the difference between baseline performance, when the target is presented alone, and the performance when the target and mask are presented together. A strong correlation between a parameter and the masking strength would suggest a critical role for this parameter for explaining how masking occurs.

2.3. Experiments

2.3.1. Para- Metacontrast Masking

2.3.1.1. Methods

In para-meta-contrast masking, we tested two statistical models, namely G and GU, as mentioned before. The G model states that the masking effect occurs due to an "interference" by the mask signal on the target signal so as to impair stimulus-encoding precision for the target. This prediction will be reflected as an increased variability in an observer's responses. On the other hand, the GU model takes an additional mechanism into account, which amounts to "a reduction of target SNR by the mask signal. According to this model, a decrease in target SNR will lead to an increase in the guess rate (modeled

by an increasing weight of the uniform distribution and a decreasing weight for the Gaussian distribution; note that these two weights add to unity).

General methods and procedures were followed. Specific to para-/meta-contrast experiment, the target luminance was either 25 cd/m^2 or 30 cd/m^2 depending on the observer. The value was chosen to yield a considerable drop in performance due to masking (at least 15% transformed performance drop from baseline). In para-meta-contrast masking, SOA values were -100, -50, -10, 0, 20, 40, 50, 60, 80, 110, 150, 200 ms. The mask was a non-overlapping ring having 1.1 deg inner and 1.4 deg outer diameters, respectively, as shown in Figure 2-1b. The luminance of the mask was 5 cd/m^2.

2.3.1.2. Results and Discussion

Mean error distributions for several SOA values are given in Figure 2-4. Distribution of response errors follows a Gaussian-like distribution at SOA values where visual masking is weak or absent (i.e., SOA < 0 ms or SOA > 60 ms). However, at SOA values where there is strong masking (e.g., SOA = 40 ms), the tails of the error distribution increases, indicating the involvement of a uniform component.

In order to quantitatively assess these qualitative observations, we fitted observers' response errors with the statistical models described before. We used Bayesian Model Comparison (BMC) to compare model performances. This method returns the average log-likelihood of each model over the selected parameter space (see Table 2-I) given the observed response errors. We then averaged the log-likelihoods across all SOAs for each observer and subtracted the average likelihood of the G model from that of GU model (see Table B-I in Appendix B for individual BMC differences and

corresponding Bayes factors). In this notation, a difference of Δ means that the observed responses are, on average, e^{Δ} times more likely under the GU model.

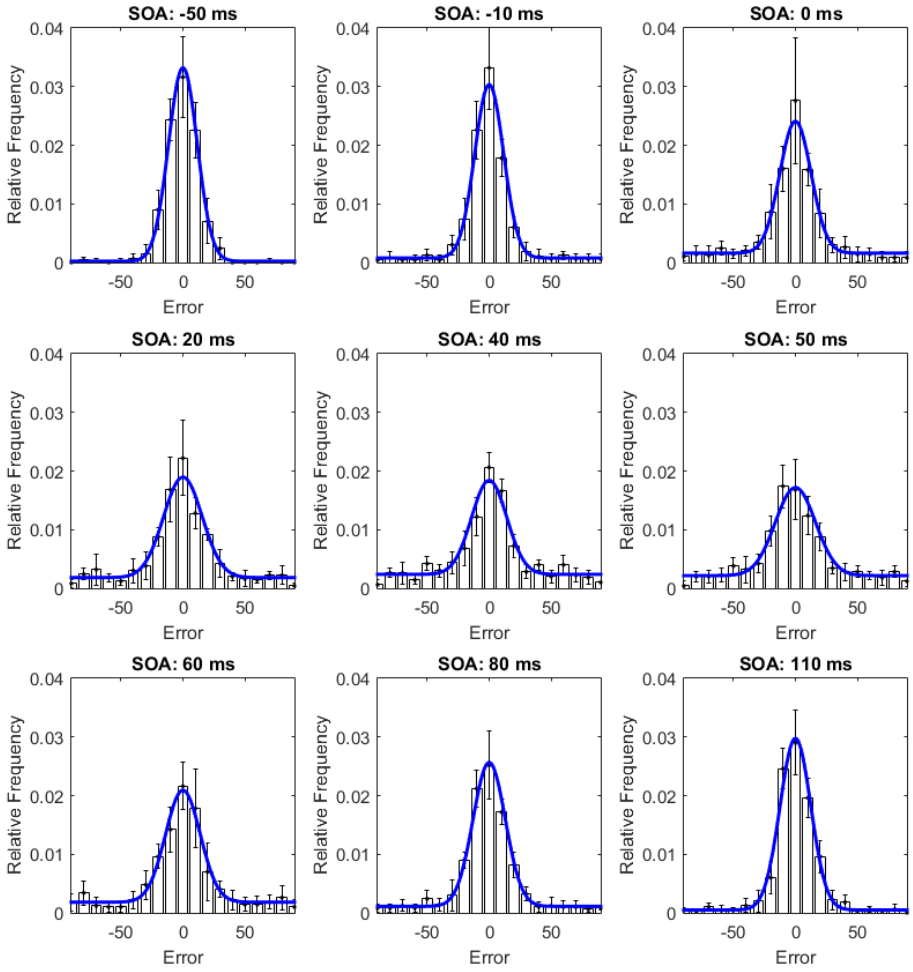

Figure 2-4 Mean error distributions and GU model fits in para/meta-contrast masking. Best fitting GU models are shown with solid blue lines. Model fits are generated by using the model parameters averaged across observers. Error bars represent SEM across observers (n=5).

Pooled across SOAs, the GU model outperformed the G model for all observers in para/meta-contrast masking experiment. Averaged across observers, log-likelihoods (i.e., BMCs) were 10.6 ± 3.9 units larger for the GU model. This corresponds to ~40,000-to-1 odds favoring the GU model. According to Jeffrey's scale of interpretation (Jeffreys,

1998), this corresponds to a “decisive evidence” for the GU model and indicates that in the present study “guessing” was an essential part of para/meta-contrast masking. Next, we extracted the model parameters of the GU model that resulted in the maximum average log-likelihood for each observer. We sought to find correlations between masking functions (more specifically, masking strengths, defined as the differences between performances with and without the masks). Figure 2-5 shows para/meta-contrast masking functions (Figure 2-5A), corresponding average masking strengths (Figure 2-5B), average model parameters for the best fitting (GU) model (Figure 2-5C and 2-5D), and the correlations of each model parameter with masking strengths.

Masking functions along with the average baseline performance of all observers are shown in Figure 2-5A. The horizontal axis shows the SOA between target and mask stimuli, whereas the vertical axis represents the transformed performance. As expected, performance shows typical Type-B U-shaped patterns with dips occurring at positive SOA values. In other words, the masking strength, defined as the drop in performance from baseline, reaches its maximum at a positive SOA (Figure 2-5B).

The standard deviation of the Gaussian in the GU model increases as SOA values approach 50ms (where masking is most effective) and then decreases to a plateau (Figure 2-5C). A one-way ANOVA confirms a significant effect of SOA on standard deviation ($F(11,44)=5.259$, $p<0.001$; Bayes factor: 618 ± 0.4%). The weight of the uniform distribution also shows a significant change with SOA ($F(11,44)=14.680$, $p<0.0001$; Bayes factor: 1.4E+9 ± 1.6%). Visual comparison of model coefficients (Figure 2-5C and 2-5D) with masking strengths (Figure 2-5B) reveals that the standard deviation of the Gaussian term and the weight of the Uniform term do correlate with masking strengths,

the latter having a stronger correlation than the former. Pearson's R coefficients confirm these qualitative observations. We found that both the standard deviation of Gaussian and the weight of Uniform in the GU model strongly correlate with the masking strength (one sample t-test results show $p<0.0001$ for both parameters).

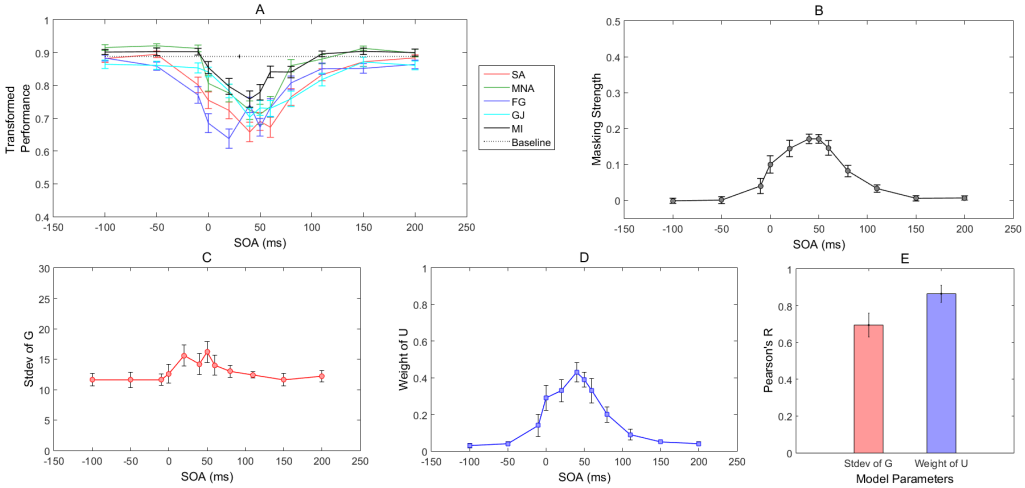

Figure 2-5 A. Para/meta-contrast masking functions for each observer. B. Mean masking strengths. Model parameters are presented here for only the winning GU model. C. The standard deviation of the Gaussian. D. The guess rate. E. Correlation of model parameters with masking strengths.

Stronger correlation between the weight of the Uniform distribution and the masking strength indicates that, as the masking strength increases, observers tend to guess more, suggesting that the target SNR is reduced by the mask activity in the visual system. Our results also suggest that the "interference" of the mask signal with the target signal, which is manifested by the increased standard deviation of the Gaussian term in the model and also by the significant correlation with the masking strength, is also present.

2.3.2. Pattern Masking by Noise

2.3.2.1. Methods

Similar to para-/meta-contrast masking, we tested two statistical models namely G and GU as mentioned before. The general methods and procedures were identical to para-/meta-contrast masking experiments. Specific to the pattern masking by noise experiment, the target luminance was 25 cd/m^2 for all observers. In pattern masking by noise, SOA values were -100, -70, -50, -30, -10, 0, 10, 30, 50, 70, 100, 150 ms. The noise mask, as shown in Figure 2-1c, was composed of 70 randomly located disks (diameters ranging from 0.2 deg to 0.3 deg), which could overlap and were confined to 2x2 deg area. We made sure that the noise masks did not have bias in any particular orientation, by obtaining 2D Fourier transforms of 100 randomly generated masks. Visual inspection of the magnitude and phase responses revealed no significant peaks or dips, hence no biases in any orientation. The luminance of the disks composing the mask was 5 cd/m^2.

2.3.2.2. Results and Discussion

Mean error distributions for several SOA values are given in Figure 2-6. Distribution of response errors follows a Gaussian-like distribution at SOA values where visual masking is weak or absent (i.e., SOA < -30 ms or SOA > 30 ms). At SOA values where masking is strongest (e.g., SOA = 0 ms), performance is near chance and the distribution of errors is uniform.

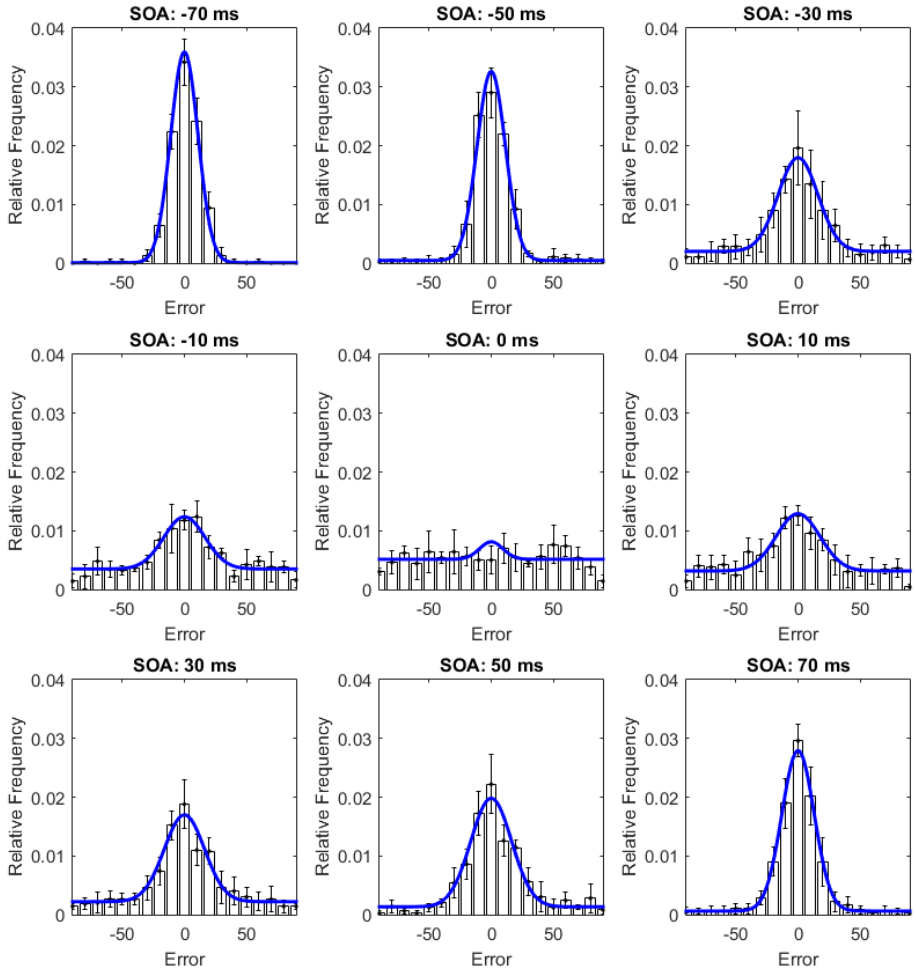

Figure 2-6 Mean error distributions and GU model fits in pattern masking by noise. Best fitting GU models are shown with solid blue lines. Model fits are generated by using the model parameters averaged across observers. Error bars represent SEM across observers (n=5).

The GU model outperformed the G model for all observers in noise masking (see Table B-I in Appendix B for individual BMC differences and corresponding Bayes factors). Averaged across observers, log-likelihoods (i.e., BMCs) were 6.4 ± 3.5 units larger for the GU model. This corresponds to ~600-to-1 odds favoring the GU model. According to Jeffrey's scale of interpretation (Jeffreys, 1998), this corresponds to "decisive evidence" for the GU model and indicates that "guessing" is an essential part of

pattern masking by noise. Individual masking functions along with the average baseline performance of all observers are shown in Figure 2-7. The horizontal axis again represents the SOA between target and mask stimuli whereas the vertical axis shows the transformed performance. As expected (Breitmeyer & Ogmen, 2006), performance shows a Type-A masking function with the strongest masking occurring at 0 ms SOA (Figure 2-7B).

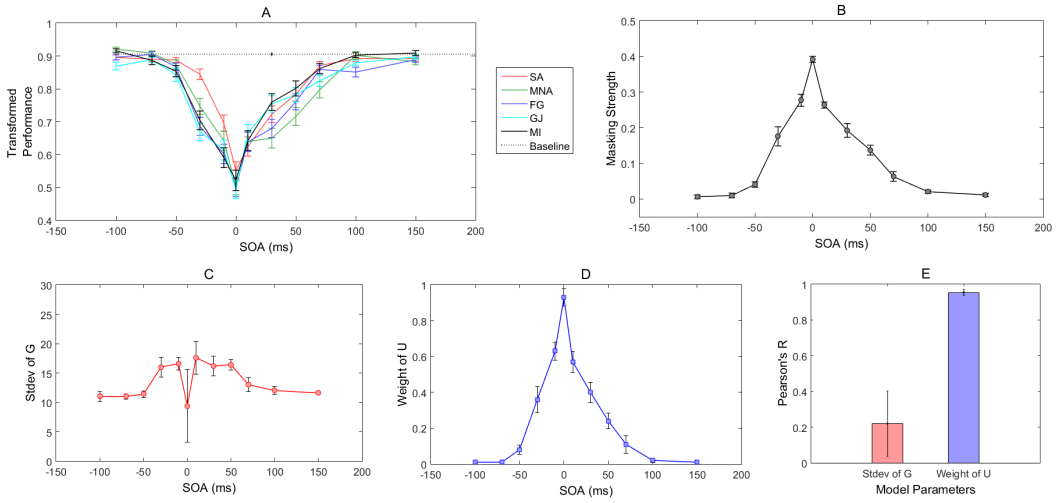

Figure 2-7 A. Masking functions in pattern masking by noise for each observer. B. Mean noise masking strengths. Model parameters are presented here for only the winning GU model. C. The standard deviation of the Gaussian. D. The guess rate. E. Correlation of model parameters with masking strengths.

Figure 2-7C and 2-7D show model parameters against SOA values. Standard deviations of the Gaussian in the GU model appears to change as a function of SOA (Figure 2-7B and 2-7C); however, a one-way ANOVA of standard deviations yielded no significant effect of SOA ($F(11,44)=1.775$, $p=0.088$; Bayes Factor: 1.2 ± 1.1%). Consistently, we found no significant correlation (average $R = 0.220$, one sample t-test: $p=0.294$) between standard deviation and masking strength. In contrast, as with para-/meta-contrast masking, guess rate strongly correlates with the masking strength (Figure

2-7B – 2-7D): The stronger the masking effect, the higher the guess rate, reflected in the weight of the uniform component in the GU model. This SOA-dependent modulation of "guessing", i.e., the weight of the Uniform, is highly significant ($F(11,44)=59.130$, $p<0.0001$; Bayes Factor: 11.5E+20 ± 2.7%). Correlation of the weights with the masking strength was also highly significant for all observers ($p<0.0001$).

In summary, these results suggest that pattern masking by noise exerts its effect mainly by reducing the SNR of the target. Since the mask consists of noise, it is reasonable to assume that SNR is reduced by increasing the noise that co-exists with the target signal.

2.3.3. Pattern Masking by Structure

2.3.3.1. Methods

In pattern masking by structure, the mask elements share structural properties of the target. Therefore, the possibility of observers' reporting one of the mask elements instead of the target stimulus cannot be discounted. For instance, the mask element that has the closest angle to the target angle, or the one that has the closest location to the target location may be reported by mistake. In this case, the masking effect, which is caused by incorrect feature binding, i.e. misattribution of orientation of a mask element to the target is modeled by an extra Gaussian distribution. In addition to the G and GU models, we tested two different misbinding models, namely Closest Angle (GUCA) and Nearest Neighbor (GUNN) for pattern masking by structure. These models have a separate Gaussian term in addition to the Gaussian and Uniform components. In the GUCA model, the mean of the second Gaussian term is determined by the mask element

that has the closest angle to the target orientation whereas it is determined by the mask element that is the nearest neighbor of the target bar in the GUNN model.

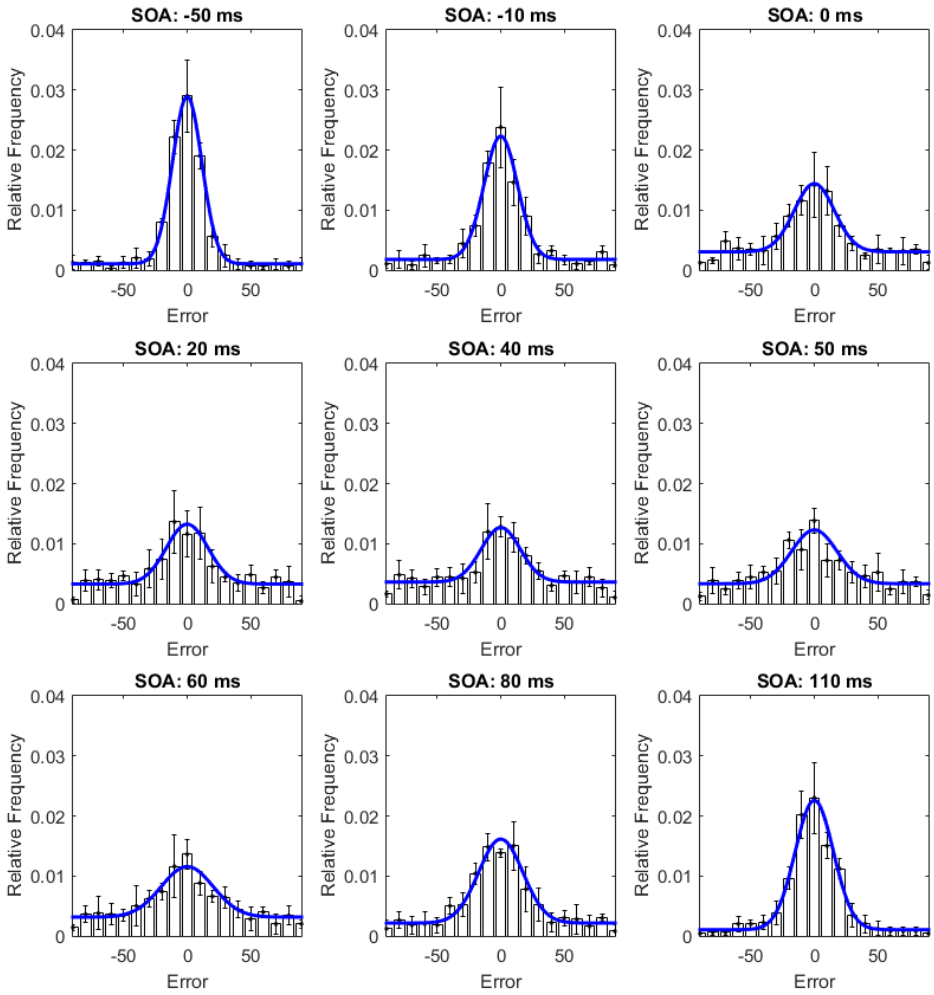

Figure 2-8 Mean error distributions and GU model fits in pattern masking by structure. Best fitting GU models are shown with solid blue lines. Model fits are generated by using the model parameters averaged across observers. Error bars represent SEM across observers (n=5).

The general methods and procedures were identical to previous experiments. Specific to the current experiment, the target luminance was 5 cd/m^2 for all observers. In this experiment, SOA values were -100, -50, -10, 0, 20, 40, 50, 60, 80, 110, 150, 200 ms. The mask shown in Figure 2-1d, was composed of 3 randomly oriented bars with the

same size as the target. Mask elements were randomly located inside a 2x2 deg virtual rectangle. The mask luminance was either 10 cd/m^2 or 20 cd/m^2 depending on the observer. The value was chosen to yield a considerable drop in performance due to masking (at least 25% transformed performance drop from baseline).

2.3.3.2. Results and Discussion

Mean error distributions in pattern masking by structure for several SOA values are given in Figure 2-8. The response errors follows a Gaussian-like distribution with varying standard deviations at all SOA values. At SOA values where masking is strongest (e.g., 10 ms < SOA < 50 ms), the increased tails of the distribution suggest the involvement of a uniform component.

Figure 2-9A plots individual masking curves and average baseline performance against SOAs. Across the entire SOA range, performance shows Type-B U-shaped patterns with dips occurring at positive SOAs. However, as expected (Breitmeyer & Ogmen, 2006), at *positive SOAs* (backward masking) the functions tended to approximate a J-shape more than a U-shape (compare results of Figure 2-9A to those of Figure 2-5A).

Figure 2-10 shows individual as well as average BMC differences for all model types (see Table B-I in Appendix B for individual BMC differences and corresponding Bayes factors). For all observers, the GUCA and GUNN models performed much better than the G model. However, the GU model, once again, outperformed all other model types. The average BMC difference between the GU and G models was 10.2 ± 3.1, which corresponds to ~27,000-to-1 odds decisively favoring the GU model. The following discussion on model parameters focuses on this model.

The standard deviation of the Gaussian term and the weight of the uniform distribution in the GU model are plotted against SOA in Figure 2-9C – 2-9D. The effect of SOA on standard deviations failed to reach significance ($F(11,44)=1.515$, $p=0.160$; Bayes factor: $0.7 \pm 0.5\%$). The weight of the uniform distribution changed significantly with SOA ($F(11,44)=18.020$, $p<0.0001$; Bayes factor: $12.4E+10 \pm 0.8\%$). We found a weak but significant correlation between the standard deviation of the Gaussian and the masking strength ($R=0.322 \pm 0.113$, one-sample t-test: $p=0.003$). On the other hand, we found a strong correlation between the weight of the uniform distribution and the masking strength ($R=0.870 \pm 0.033$, one-sample t-test: $p<0.0001$). Hence, it should be noted that a major factor in producing a masking effect is a reduction in SNR because the weight of the Uniform term strictly follows the masking strength whereas the standard deviation of the Gaussian does not. Therefore, these findings suggest that pattern masking by structure also occurs, from a statistical point of view, primarily due to the reduction of target SNR and only partly, if at all, to the interference of the target signal with the mask-related activity.

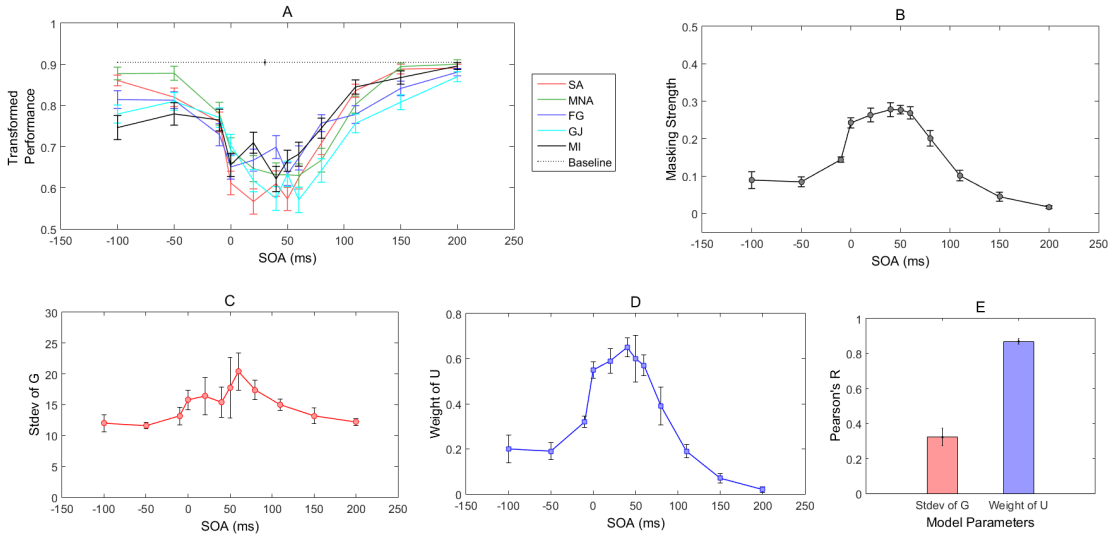

Figure 2-9 A. Masking functions in pattern masking by structure for each observer. B. Mean noise masking strengths. Model parameters are presented here for the GU model. C. The standard deviation of the Gaussian. D. The guess rate. E. Correlation of model parameters with masking strengths.

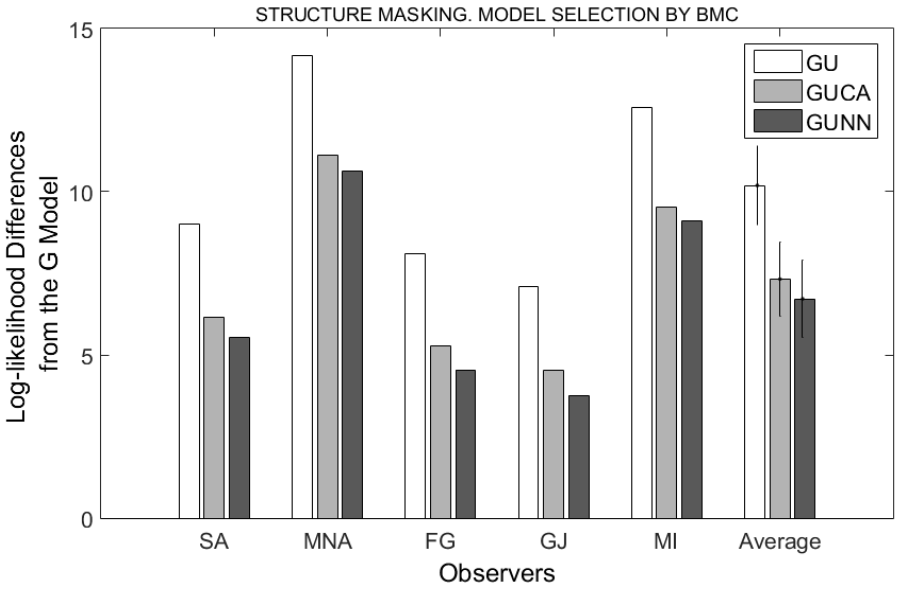

Figure 2-10 Average log-likelihood (BMC) differences between the models tested for each observer. All differences are computed by subtracting the BMC of the G model from all other model types.

2.4. General Discussion

In this study, we adopted a statistical point of view to investigate interactions between target- and mask-related activities within the context of visual masking. We modeled the distribution of the response errors of human observers in three different visual masking experiments, namely para-/meta-contrast masking, pattern masking by noise, and pattern masking by structure. Table 2-II summarizes the results. In all masking types, the GU model was the winning model, showing that a single statistical model was able to capture the response characteristics of the observers in three different masking types. We now discuss how one can interpret the parameters of the statistical model in terms of underlying mechanisms of masking.

Table 2-II Summary of the results. The results about the mean of the Gaussian are taken from Appendix A.

Masking type	Para-/meta-contrast			Masking by noise			Masking by structure		
Winning model	GU			GU			GU		
Parameters	μ_g	σ_g	w_u	μ_g	σ_g	w_u	μ_g	σ_g	w_u
Dependence on SOA	×	✓	✓	×	×	✓	×	×	✓
Correlation with masking strength	o	++	+++	o	o	+++	o	+	+++
Interpretation		signal interference	SNR ↓			SNR ↓		signal interference	SNR ↓

μ_g: mean of gaussian, σ_g: standard deviation of gaussian, w_u: weight of uniform (guess rate)

o : no correlation, + : weak positive correlation, ++ : positive correlation, +++ : strong positive correlation

In our LMS + adjusted R^2 based analysis, we did not find any systematic change in the means of the Gaussian term in the model (Appendix A), indicating no bias for any target orientation in any of the masking types. Hence, the mask did not interfere with the accuracy of target encoding. Whereas bias was not found in the masking experiments here, it is worth discussing the implications of finding a bias as we would expect biases in

two masking related phenomena. First, in feature attribution or feature inheritance, features of the target are transposed to the mask (Ağaoğlu, Herzog, & Ogmen, 2012; Enns, 2002; Herzog & Koch, 2001; Hofer, Walder, & Groner, 1989; Ogmen, Otto, & Herzog, 2006; Otto, Ogmen, & Herzog, 2006, 2008; Stewart & Purcell, 1970; Werner, 1935; A. E. Wilson & Johnson, 1985). Hence, if observers were to report features of the mask, one would expect systematic biases that are congruent with the features of the target. Second, in masked-priming studies (Ansorge, Klotz, & Neumann, 1998; Klotz & Neumann, 1999; Schmidt, 2002; Vorberg, Mattler, Heinecke, Schmidt, & Schwarzbach, 2003), observers are asked to make speeded responses to the mask (rather than the target). The congruence of a target's features with those of the mask produces faster responses. Hence, again one may expect target-congruent biases if observers were to report features of the mask. Biases would be expected as well if this modeling were to be applied to other visual phenomena, such as visual after-effects and crowding (Ester et al., 2015; van den Berg, Johnson, et al., 2012). However, the finding that the target orientations were reported without any bias in the present study was an expected consequence of the way we designed our mask stimuli, in that they did not have any systematic orientation bias.

When masking is very strong, as we observed here in the case of SOA = 0 ms in structure masking by noise, observers guess and the finding that the error distribution is uniform becomes trivial. However, there is no a priori reason to expect that, when masking strength is reduced, a uniform distribution will play a major role in explaining masking. In fact, due to the choice of stimulus parameters used in this study, observers are at chance only for SOA = 0 ms in pattern masking by noise, but not in para-/meta-

contrast masking and pattern masking by structure. Therefore, the fundamental role identified for the uniform distribution in this study is supported by the analysis of all the data points that are above chance level.

In all masking types studied here, an increase in the weight of the Uniform distribution (and equivalently a decrease in the weight of the Gaussian term) correlated most strongly with masking strengths (see Table 2-II). Since the changes in the weights of the Gaussian and Uniform terms are interpreted as changes in the target SNR, the masking effects mainly manifest themselves as a reduction of target SNR. While in the statistical model, decreases in signal strength and increases in noise are intertwined, we can speculate on the individual changes in signal strength and in noise based on the assumption that noise is most effective when it is integrated with the target signal. Accordingly, an increase in the noise component of SNR would be most effective at 0 ms SOA when target and mask temporal integration is maximal. Based on this assumption we suggest that:

1. In metacontrast, relatively weak masking occurs at 0 ms SOA and maximum masking occurs at positive SOA (U-shaped Type-B), implying that masking occurs mainly by the reduction of the *signal* in the SNR, with the mask interrupting or suppressing the strength of the target activity.
2. In pattern masking by noise, maximum masking occurs at 0 ms SOA (Type-A), implying that masking occurs mainly by an increase in the *noise* component of the SNR.

3. In pattern masking by structure, one obtains strong masking at 0 ms SOA and maximum masking at positive SOA (J-shaped Type-B), implying that masking occurs both by increases in noise and decreases in signal of the SNR.

Chapter 3. Interactions between Spatial Attention and Visual Masking[5]

3.1. Introduction

Visual masking is defined as the reduction of visibility of one stimulus (target) by another stimulus (mask) when the mask is presented in the spatio-temporal vicinity of the target (Bachman, 1994; Breitmeyer & Ogmen, 2006). Visual masking has largely been investigated as a phenomenon reflecting the spatiotemporal dynamics of the visual system, and it provides a useful tool to study differences between nonconscious stimulus- and conscious percept-dependent visual processing. Several types of masking depending on the spatiotemporal characteristics of the stimuli have been identified. When the target is followed by the mask in time, it is referred to as *backward masking* whereas when the mask precedes the target, it is called *forward masking*. Moreover, when the target and mask onsets coincide but the mask outlasts the target, it is called *common-onset masking*. In terms of spatial properties, backward masking is referred to as *metacontrast masking* when the target and mask stimuli do not spatially overlap.

In terms of information processing, it is known that visual masks can suppress, or "erase," the contents of sensory (or iconic) memory, which is a large capacity and rapidly decaying store (Averbach & Sperling, 1961; Haber, 1983; Sperling, 1960). The control of the contents of sensory memory by masking mechanisms has two important functional implications: First, since the contents of sensory memory are encoded in retinotopic

[5] All of the finding reported here have recently been submitted for publication.

coordinates, based on the duration of the visible-persistence component of sensory memory, moving objects should appear highly smeared. Empirical and computational evidence shows that, by suppressing the contents of sensory memory, visual masking mechanisms play an important role in establishing the clarity of our vision for moving objects (Chen et al., 1995; Noory, Herzog, & Ogmen, 2015; Ogmen, 1993; Purushothaman et al., 1998). Second, a subset of the contents of sensory memory is transferred to a more durable but low-capacity store, called visual short-term memory (VSTM) (Atkinson & Shiffrin, 1971; Averbach & Sperling, 1961). One of the distinguishing properties of VSTM from sensory memory is its immunity to visual masking (e.g., Averbach & Coriell, 1961; Gegenfurtner & Sperling, 1993; Haber, 1983; Loftus, Duncan, & Gehrig, 1992; Schill & Zetzsche, 1995). Hence, visual masking plays an important functional role in controlling which information will be available for transfer to VSTM.

Another process known to control the transfer of information from sensory memory to VSTM is attention (e.g., Gegenfurtner & Sperling, 1993; Makovski & Jiang, 2007; Ogmen, Ekiz, Huynh, Bedell, & Tripathy, 2013; Palmer, 1990; Sreenivasan & Jha, 2007; Tombu et al., 2011). Since both attention and visual masking control (i.e., modulate) the transfer of information from sensory memory to VSTM, a natural question is whether these processes operate independently or they interact with each other. From a theoretical point of view, determining whether these two processes interact or not can contribute to our understanding of how information is transferred from sensory memory to VSTM. From an empirical point of view, this understanding is especially important when one wants to compare findings from different studies of VSTM, which employ

different types of masks or masks with different strengths. If, indeed, masking and attention interact, reconciliation or comparison of findings across different studies will require one to take into account the interaction effects.

Determining whether masking and attention interact also has important implications on theories of visual masking. Selective attention has facilitative, as well as inhibitory, effects in almost all perceptual tasks and regardless of criterion contents (Posner, 1980; Smith et al., 2004). However, many early theoretical models of masking do not include a term or a mechanism for the effects of attention, implying that these models assume that attention and masking are independent processes (e.g., Bachmann, 1984; Bruno G Breitmeyer & Ganz, 1976; Bridgeman, 1971; Francis, 2000; Ogmen, 1993; Weisstein, Ozog, & Szoc, 1975). This does not necessarily mean that these models dismiss the role of attention. Attention can be incorporated to these models largely as an add-on process, which adds to the masking strength, or reduces it, depending on the locus of attention or attentional load. In fact, Michaels and Turvey (1979) incorporated attention in their model as an independent process working in conjunction with spatial inhibitory processes.

On the other hand, at least one theory of visual masking considers attention as an essential component and predicts interactions between masking and attention (Di Lollo et al., 2000; Enns & Di Lollo, 1997). In a common onset masking paradigm, Enns and Di Lollo (1997) used a diamond shaped stimulus as target and four surrounding dots as mask. They found that the four-dot mask can produce strong masking effects when the stimuli were viewed peripherally *and* when attention could not be focused on a certain target location (i.e., with set sizes larger than one). Enns and Di Lollo attributed these

effects to higher-level processes of *object substitution*. Here, the assumption is that, as set-size increases, attentional resources will have to be spread over more locations, thereby increasing the attentional load and, hence, the time it takes for focused attention to arrive at the target's location. On the other hand, when set-size is small or when the target just "pops out", attention quickly focuses on this location. If attention arrives to the location of the target before re-entrant signals feedback to the target's location, the observer will be able to perceive and identify the target. On the other hand, if re-entrant signals arrive at the target's location before attention, a mismatch between the reentrant visual representation of the target-mask pair and the incoming lower level activity due to mask alone (since it is presented alone after target's offset) will occur. In this case, the mask-only representation will substitute in perception the early activities generated by the target-mask pair. In summary, interaction between attention and masking is an essential ingredient of the object substitution theory. This prediction was supported by significant interaction effects found in their study (Di Lollo et al., 2000; Enns & Di Lollo, 1997).

Reports of interactions between masking and attention have not been limited to the common-onset masking paradigm, but also included metacontrast masking (Ramachandran & Cobb, 1995; Shelley-Tremblay & Mack, 1999; Tata, 2002). Hence, a question arises as to whether *all* theories of masking should include attention as an essential component.

However, more recent evidence shows that, in common-onset masking with four-dot masks, masking strength and set size (i.e., attention), do not actually interact, and that previous studies suffered from ceiling and/or floor effects which led to artifactual appearance of interactions (Argyropoulos et al., 2013). This finding has been recently

replicated by using an eight-alternative forced-choice task (Filmer et al., 2014), providing further evidence against the attention account of object-substitution theory. Pilling et al. (2014) also employed a spatial cue to directly control spatial attention, and also reported no interaction. Filmer, Mattingley, and Dux (2015) have also demonstrated strong common-onset masking for the attended and *foveated* targets, which strongly contradicts the object-substitution account of common-onset masking. Given these findings, we have examined whether the reported interactions between attention and metacontrast may also be artifacts of ceiling and /or floor effects. The objective of the present study was to investigate whether metacontrast masking and attention interact by using an experimental design in which saturation and floor effects are avoided. We asked observers to report the orientation of a target bar when presented with other randomly tilted distractor bars. By adjusting stimulus parameters for each observer such that both the ceiling and floor effect are avoided, we investigated the relationship between masking and attention at two different levels: (i) in mean absolute response errors and (ii) in distribution of signed response errors. Our results show that although attention affects observer's performance, its effect does not interact with masking. Statistical modeling of response errors suggest that attention and masking exert their effects by independently modulating the probability of "guessing" behavior.

3.2. Methods

3.2.1. Participants

Seven observers participated in this study (three female, four male; ages from 24-28). Five of them were naïve as to the purpose of the experiment. Participants reported

normal or corrected-to-normal vision and gave written informed consent before the experiments. All experiments were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and followed a protocol approved by the University of Houston Committee for the Protection of Human Subjects.

3.2.2. Apparatus

Visual stimuli were created using the ViSaGe and VSG2/5 cards manufactured by Cambridge Research Systems. Stimuli were displayed on a 22-in. CRT monitor with a refresh rate of 100 Hz and display resolution of 800 by 600 pixels. The distance between the display and the observer was 1 m, and a head/chin rest was utilized to restrict movements of the observer. Observers responded via a joystick after each trial.

3.2.3. Stimuli and Procedures

In each trial, several oriented bars equidistant from the display center were presented briefly (10 ms). Any one of the bars could potentially be the target stimulus, and the target was specified by the mask location. In other words, only one mask stimulus was presented and its location cued which oriented bar is the target. The other bars will be referred to as distractors from now on. The task of the observers was to report the orientation of the target bar. Figure 3-1 illustrates the stimuli and procedures. A trial starts with a black (0 cd/ m^2) fixation spot at the center of a blank white screen (60 cd/m^2). After a random time interval (500-1000 ms), an array of randomly oriented bars at equal eccentricities was presented around an imaginary circle (with a radius of 6 deg). After an SOA interval (0 – 200 ms), a spatially non-overlapping mask stimulus (a ring) was presented. In separate blocks, a small square (rather than a mask) indicating the

location of the target bar was presented in some of the trials. The trials where a small square was presented were considered as the baseline trials. The duration of the target and mask was 10 ms (one frame). The luminance of the target and mask was adjusted individually for each observer to avoid floor and ceiling effects. Once the target-mask sequence was presented and turned off, a randomly oriented (response) bar was displayed at the center of the screen, and observers adjusted its orientation (illustrated by red arrows in Figure 3-1) via a joystick to match the target bar's orientation. The response bar stayed on the display until observers were satisfied with their responses and the next trial began with another button press. In separate block of trials, we presented an array with two or six oriented bars. Varying the set size allowed us to determine the effect of attention and its interaction, if any, with masking.

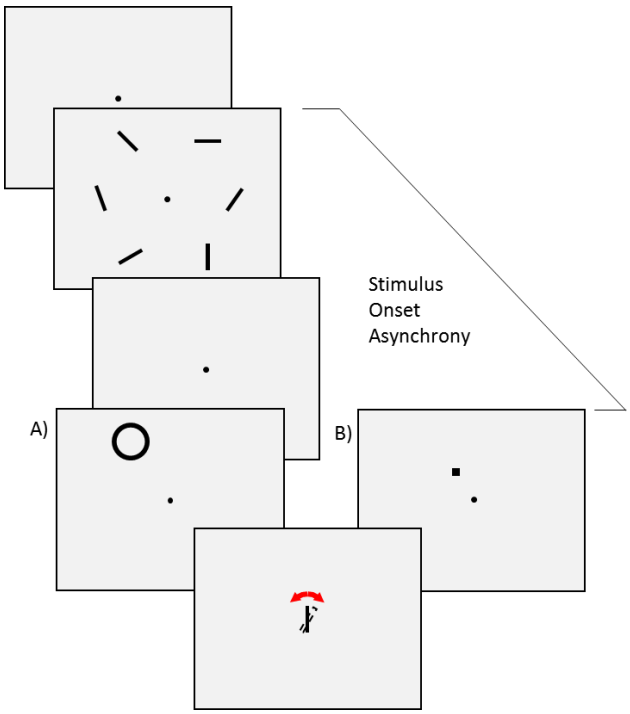

Figure 3-1 Time course of the stimuli. The target array was followed either by (A) a mask stimulus (a ring), or by (B) a small square (a post-cue) which were presented for 10 ms. The task of the observers was to report the orientation of the masked (or probed) bar.

We defined response errors as the difference between the actual and reported orientations. Error values ranged from -90 to 90 deg. We obtained masking functions after transforming response errors to a probability like measure such that performance values of 0.5 and 1 correspond to chance and perfect performance, respectively. We calculated transformed performance (Ogmen et al., 2013) as

$$Transformed\ Performance = 1 - \frac{|Error\ Angle|}{90} \quad . \tag{3-1}$$

When the observer can report the orientation of the target bar veridically, then error angle will be zero, which corresponds to a transformed performance value of 1. When the observer randomly guesses, the absolute value of the response error will be distributed uniformly within the range of 0 and 90 deg. Hence, the average of the absolute value of error angles will be 45 deg with the corresponding Transformed Performance value equal to 0.5.

For the purpose of this study, ceiling and floor effects must be avoided. The floor can be defined as the chance level, which corresponds to 0.5 transformed performance (see Equation 3-1), and the ceiling can be defined as the maximum performance an observer can possibly achieve in the absence of a mask. Thus, one has to determine the ceiling level (i.e., baseline performance in the absence of a mask) for each and every observer by presenting the target stimulus only. However, since we present an array of oriented bars rather than a single one, one needs to specify which one of them is the target without affecting its visibility. Moreover, obtaining the baseline performance at a single SOA may not be appropriate, since there may be additional confounding factors (such as memory leakage especially at long SOAs). Therefore, we presented a small

square (0.2 deg x 0.2 deg) as a cue in the spatial vicinity of the target bar at each SOA. This way we had a separate baseline performance for each SOA, and we made sure that the ceiling effect is avoided at each SOA.

Before the actual experiments, we first trained the observers with two or three blocks of trials with all conditions to make sure that they became familiar with the experiment and the setup so as to stabilize their performance and minimize changes due to learning. In order to avoid the ceiling and floor effects described above, we adjusted two parameters: the target luminance and the mask luminance. The criteria that we used to obtain the target and mask luminance values were as follows:

C1) The maximum performance with masking must be significantly lower than the baseline performance (the ceiling) when set size is two.

C2) The minimum performance with masking must be significantly higher than chance level (the floor, i.e., 0.5 transformed performance).

Based on pilot experiments and our previous studies on metacontrast masking, we carried out a power analysis to select the number of trials per SOA for masking and baseline (i.e., without a mask) conditions. Power analysis is necessary to assess Type-I errors (i.e., probability of falsely rejecting a null hypothesis). Therefore, we determined the number of trials required to reject the null hypothesis (i.e., the transformed performances with and without a mask are equal) by a two-sample t-test with a power level larger than 0.7. This analysis yielded roughly 200 trials per SOA value in total. Therefore, each observer ran 125 masking trials and 75 baseline trials per SOA. Table 3-I summarizes the target-mask luminance pairs for all observers as well as the results of statistical tests to verify that the aforementioned criteria (C1 and C2) are met.

Table 3-I The target, mask, and cue luminance values in cd/m^2 (and Weber contrasts) for each observer. The background was set to 60 cd/m^2 for all observers. The results of t-tests used to determine whether criteria C1 and C2 are met are also listed.

	Luminance (Contrast)			Statistical Criteria	
Observer	**Target**	**Mask**	**Cue**	**C1 (ceiling)**	**C2 (floor)**
AK	*39.5 (-0.34)*	*6 (-0.9)*	*30 (-0.50)*	t(190.3) = -1.68; p = 0.047	t(124) = 3.96; p < 0.001
CBK	*20 (-0.67)*	*0 (-1.00)*	*21 (-0.65)*	t(183.1) = -2.97; p = 0.002	t(124) = 4.34; p < 0.001
EK	*48 (-0.20)*	*0 (-1.00)*	*48 (-0.20)*	t(190.3) = -1.86; p = 0.032	t(124) = 4.05; p < 0.001
FG	*49.5 (-0.18)*	*6.5 (-0.89)*	*17 (-0.72)*	t(192.5) = -2.18; p = 0.015	t(124) = 3.74; p < 0.001
GQ	*42 (-0.30)*	*10 (-0.83)*	*22 (-0.63)*	t(181) = -3.16; p < 0.001	t(124) = 3.3; p < 0.001
MNA	*42 (-0.30)*	*15 (-0.75)*	*20 (-0.67)*	t(153.4) = -4.49; p < 0.001	t(124) = 4.71; p < 0.001
SA	*47.5 (-0.21)*	*0 (-1.00)*	*25 (-0.58)*	t(193.2) = -3; p = 0.002	t(124) = 2.26; p = 0.013

Since masking strength is observer-dependent, the same set of parameters for all observers may not avoid floor and ceiling effects. For this reason, we adjusted target and mask luminance values individually for each observer to make sure that the data were free of floor and ceiling effects. Moreover, changing target and mask luminance does alter the location of maximum masking and may even result in Type-A (i.e., maximum masking at 0 ms SOA and monotonic increase in performance for positive SOAs) or Type B (i.e., maximum masking at a positive SOA and minimal or no masking at 0 ms SOA and beyond 300 ms SOA) masking functions depending on the observer. Therefore, the luminance values should be adjusted for each observer separately in order to produce Type-B masking functions, a prominent signature of metacontrast masking, for each observer. In order to capture the "U-shaped" masking functions from each and every observer, we needed to select a different set of SOA values.

Table 3-II The regression models used to fit transformed performances and the winning model parameters. The models are sorted based on number of parameters.

ID	*Regression Model*
M1	$Y = \beta_0 + \varepsilon$
M2	$Y = \beta_0 + \beta_1 \tau + \varepsilon$
M3	$Y = \beta_0 + \beta_1 n + \varepsilon$
M4	$Y = \beta_0 + \beta_1 \tau n + \varepsilon$
M5	$Y = \beta_0 + \beta_1 \tau^2 + \varepsilon$
M6	$Y = \beta_0 + \beta_1 \tau^2 n + \varepsilon$
M7	$Y = \beta_0 + \beta_1 \tau + \beta_2 n + \varepsilon$
M8	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau n + \varepsilon$
M9	$Y = \beta_0 + \beta_1 n + \beta_2 \tau n + \varepsilon$
M10	$Y = \beta_0 + \beta_1 \tau^2 + \beta_2 n + \varepsilon$
M11	$Y = \beta_0 + \beta_1 \tau^2 + \beta_2 \tau^2 n + \varepsilon$
M12	$Y = \beta_0 + \beta_1 n + \beta_2 \tau^2 n + \varepsilon$
M13	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \varepsilon$
M14	$Y = \beta_0 + \beta_1 \tau + \beta_2 n + \beta_3 \tau n + \varepsilon$
M15	$Y = \beta_0 + \beta_1 \tau^2 + \beta_2 n + \beta_3 \tau^2 n + \varepsilon$
M16	$\mathbf{Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \beta_3 n + \varepsilon}$
M17	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \beta_3 \tau n + \varepsilon$
M18	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \beta_3 \tau^2 n + \varepsilon$
M19	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \beta_3 n + \beta_4 \tau n + \varepsilon$
M20	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \beta_3 n + \beta_4 \tau^2 n + \varepsilon$
M21	$Y = \beta_0 + \beta_1 \tau + \beta_2 \tau^2 + \beta_3 n + \beta_4 \tau n + \beta_5 \tau^2 n + \varepsilon$

τ : SOA, n : Set size, ε : Error term.

Once we established a set of parameters which satisfied all the criteria given above, we analyzed transformed performance of each observer separately (within-subject analysis). We fitted a series of linear and polynomial regression models to pin down the presence/absence of contributions of the main factors and their interactions. Table 3-II

lists all regression models used to fit the data. We used Bayesian Information Criterion (BIC) and Adjusted R^2 metrics for selecting the best model. Both metrics resulted in similar, if not identical, model selections for all observers. Both metrics penalize the models with more free parameters. Absolute values of BICs are not meaningful, therefore one needs to look at differences between BICs from different models. A BIC difference of x between model A and model B (i.e., $x = BIC_A - BIC_B$) corresponds to e^{-x} –to-1 odds favoring model A. Therefore, the smaller the BIC, the better the model performs. According to Jeffreys' scale of interpretation (Jeffreys, 1998), an odds ratio lower than one (i.e., $e^{-x} < 1$) supports the null hypothesis, whereas an odds ratio larger than one (i.e., $e^{-x} > 1$) supports the alternative hypothesis. Values larger than 100 ($e^{-x} > 100$) are considered as a sign of "decisive evidence" against the null hypothesis, and similarly, values smaller than 0.01 (i.e., $e^{-x} < 0.01$) are interpreted as "strong evidence" against the alternative hypothesis.

3.2.4. Statistical Modeling of Response Errors

We examined the distribution of response errors of observers to understand how attention and masking exert their effect on performance. We adopted the statistical models that have been previously used in modeling VSTM (Bays et al., 2009; Zhang & Luck, 2008) and several visual phenomena such as crowding (Ester et al., 2015) and masking (Agaoglu, Agaoglu, Breitmeyer, & Ogmen, 2015; Harrison, Rajsic, & Wilson, 2014). The simplest model is a single Gaussian (referred to as the G model from now on) whose mean and standard deviation may be modulated by attention and/or masking. The mean of Gaussian represents how accurately the target orientation is encoded by the visual system. Nonzero values indicate observer bias in responses. The mean of the target

Gaussian was set to zero, i.e., centered on target orientation. This was motivated by our recent study on masking where we found that the mean of the Gaussian is not significantly different from zero (Agaoglu et al., 2015). Therefore, in the following analyses, the target Gaussians were centered on target orientations (i.e., zero mean in error space). The reciprocal of standard deviation represents how precisely the stimulus falling onto retina is encoded by the visual system. In other words, decreased stimulus encoding precision is reflected by the increased variability of behavioral responses.

In the second model, Gaussian + Uniform (the GU model), the additional Uniform component represents the "guess rate". The increased guess rate is modeled by the weight of the uniform distribution. The GU model is a weighted sum of Gaussian and Uniform distributions (Equation 3-2),

$$PDF(\varepsilon) = w_G * G(\mu, \sigma) + (1 - w_G) * U, \quad (3\text{-}2)$$

where $PDF(\varepsilon)$ represents the distribution of response errors, and w_G represents the weight of the Gaussian term with mean and standard deviation given as μ and σ.

Since the stimulus display consists of multiple oriented bars, observers may report the orientation of one of the non-target bars, e.g., the one that has the closest angle to the target angle (the GU + Closest Angle model or in short, the GUCA model), or the closest location to the target location (the GU + Nearest Neighbor or in short, the GUNN model), instead of the target bar. The contribution of this incorrect identity binding errors can be captured by another Gaussian term in the model. In consequence, the PDF of response errors can be written as a weighted sum of the target Gaussian, non-target Gaussian, and a uniform component (Equation 3-3) (Bays et al., 2009).

$$PDF(\varepsilon) = w_T G(\mu_T, \sigma_T) + w_{NT} G(\mu_{NT}, \sigma_{NT}) + (1 - w_T - w_{NT}) U(-\pi/2, \pi/2), \quad (3\text{-}3)$$

where subscripts T and NT represent target and non-target parameters, respectively.

3.2.5. Model Fitting and Model Comparison

We used the Bayesian Model Comparison (BMC) technique (Mackay, 2004; Wasserman, 2000) for selecting the best fitting model. Each model m_j produces a conditional probability $p(\varepsilon|m_j, \theta)$, where ε is a vector of observed response errors, and θ is a vector of model parameters. For each model, we calculated the probability of finding observed response errors, averaged over free parameters:

$$L(m_j) \triangleq p(\varepsilon/m_j) = \int p(\varepsilon|m_j, \theta) p(\theta|m_j) d\theta = \int \left(\prod_{i=1}^{N} p\left(\varepsilon_i \middle| m_j, \theta\right)\right) p\left(\theta \middle| m_j\right) d\theta, \quad (3\text{-}4)$$

where N represents the number of trials and ε_i represents the error in the i^{th} trial. It is convenient to take the logarithm of Equation (3-4) in order to compute it numerically. Equation (3-4) can be rewritten as

$$\ln L(m_j) = \ln L_{max}(m_j) + \ln[\int exp(\ln L(m_j|\theta) - \ln L_{max}(m_j)) \, p\left(\theta \middle| m_j\right) d\theta], \quad (3\text{-}5)$$

where $\ln L(m_j|\theta) = \sum_{i=1}^{N} ln\,(p\left(\varepsilon_i \middle| m_j, \theta\right)$, and $L_{max}(m_j) = \max(L(m_j|\theta))$. Parameters corresponding to $L_{max}(m_j)$ can be regarded as the Maximum Likelihood Estimation (MLE) of the model parameters for model m_j. Subtracting $L_{max}(m_j)$ ensures that the exponential in the integrand is of order 1 and thereby, avoids numerical problems (Ester et al., 2015; Mackay, 2004; van den Berg, Shin, et al., 2012). We used uniform priors for all parameters over plausible ranges (see Table 3-II).

Considering these priors, Equation (3-5) becomes

$$\ln L(m) = \ln L_{max}(m) - \sum_j^k \ln (R_j) + \ln[\int exp(\ln L(m|\theta) - \ln L_{max}(m))\, d\theta], \quad (3\text{-}6)$$

where R_j represents the size of the range for jth free parameter. We approximated the integral by a Riemann sum with at least 50 bins in each parameter dimension (see Table 3-III). The performance metric given in Equation (3-6) will be referred to as the BMC. The difference between the BMCs from two different models is equivalent to the logarithm of likelihood ratios of them. Therefore, a model with larger BMC is a better model. We used Jeffrey's scale of interpretation for comparing BMCs from different models. A BMC difference of x between model A and model B corresponds to e^x-to-1 odds favoring model A.

Table 3-III Range of parameters used for BMC. We repeated the analysis with step sizes of 0.1 for standard deviation of the Gaussian, and 0.002 for the weight of the Uniform for the GU model but the winning model and the model parameters were not affected by this change.

	σ_T	w_U	w_{NT}	σ_{NT}
Upper bound	50	1	1	50
Lower bound	1	0	0	1
Step size	0.25	0.02	0.02	0.25

σ_T Standard deviation of the target Gaussian
w_U Weight of the Uniform
w_{NT} Weight of the non-target Gaussian
σ_{NT} Standard deviation of the non-target Gaussian

3.2.6. Analysis of Model Parameters

After selecting the best fitting model, we sought to find how different model parameters change with SOA and set size. The motivation behind this analysis was to understand whether and how masking and attention affect the statistics of observer responses. After determining the winning model, we created 500 different data sets (for each observer separately) by resampling the response errors by replacement, and fitted

the winning model to these data sets. We present here the means and standard errors for model parameters obtained from this bootstrap analysis. Next, we fitted the regression models listed in Table 3-II to see the contributions of SOA, set size, and their interactions to model parameters.

In order to determine whether masking strength and different model parameters are related or not, we also quantified the correlation between model parameters and masking function for each set-size by calculating Pearson R coefficients. A strong correlation would suggest a critical role for that parameter in accounting for masking effects, and a change in correlation with set size would suggest an interaction between attention and masking.

3.3. Results

3.3.1. Psychophysics

Figure 3-2 shows results from all observers. The vertical axes represent the transformed performance while the horizontal axes represent SOA between the target and mask (or cue in the baseline conditions) stimuli. Open and filled symbols correspond to the baseline and masking conditions, respectively. Circles and squares plot the results for set-size of 2 and 6, respectively. Consider first the baseline data[6]. Our goal in collecting the baseline data was to ensure that the masking data did not have any ceiling effect (criterion C1, see Methods). For each observer, we performed a two-sample t-test between baseline and masking conditions at an SOA value where minimum masking occurs (e.g., typically 0 ms SOA with set size 2). For all observers, transformed

[6] See Appendix for the regression analysis of the baseline data.

performance was significantly smaller in masking condition ($p<0.05$). In addition to ceiling effects, we have also checked our data for floor effects (criterion C2, see Methods) by performing a one-sample t-test against the chance level (i.e., 0.5 transformed performance) at SOA values where masking is strongest. We confirmed that transformed performance was significantly larger than the chance level ($p<0.05$) for all observers even when masking is strongest. Table 3-I lists the target-mask luminance pairs which allowed us to avoid ceiling and floor artefacts for each observer, as well as the results of the t-tests. Taken together, these results show that our masking data are free of ceiling and floor effects.

3.3.2. Do Attention and Masking Interact?

We fitted a series of polynomial regression models (in addition to the standard linear regression models) to each observer's data to determine whether SOA and set size and their interaction have any significant contribution to transformed performance. Figure 3-2 (the right column) shows pairwise model comparison results based on the BIC metric. Greenish colors represent equivalent performances whereas blue and red colors represent better and worse model performances, respectively. As evident from Figure 3-2, the models with quadratic and linear SOA terms perform better than any of the standard regression models. This is to be expected since type-B functions' U-shape is better captured by a quadratic term than a linear term. The key aspect of this analysis was to determine whether models with interaction terms would perform better than those without interaction terms. The model M16 was the best model for each and every observer who participated in the present study. This model consists of linear SOA and set size terms as well as a quadratic SOA term but does not have any interaction term.

Therefore, our analysis indicates that SOA (i.e., masking) and set size (i.e., attentional load) do not interact.

3.3.3. Modeling

Next, we examined the distribution of response errors of each observer by using the BMC technique (see Methods). Figure 3-3 (the leftmost column) shows BMC differences between every combination of model pairs for each observer. Among the four models tested, the GU model was the winning model for all observers; it has the highest BMC value. Averaged across observers, the BMC of the GU model was 27.2, 2.9, and 3.4 larger than the G, GUCA, and GUNN models, respectively. These differences correspond to ~6E+11-to-1, ~18-to-1, and ~30-to-1 odds, all favoring the GU model. According to Jeffreys' scale of interpretation (Jeffreys, 1998), these odds correspond to "decisive evidence" favoring the GU model. Therefore, further analyses were done on model parameters of the GU model.

Figure 3-3 also shows the model parameters for the winning GU model for all observers (the second and third columns). There is no discernable pattern that is consistent across all observers in the dependence of standard deviations on SOA and set-size (Figure 3-3, the second column). On the other hand, the weight of the uniform component has clear and consistent pattern in all observers (Figure 3-3, the third column). The weight parameter changes as a function of SOA following an inverse-U function, which reflects the shape of Type-B metacontrast functions. These inverse-U

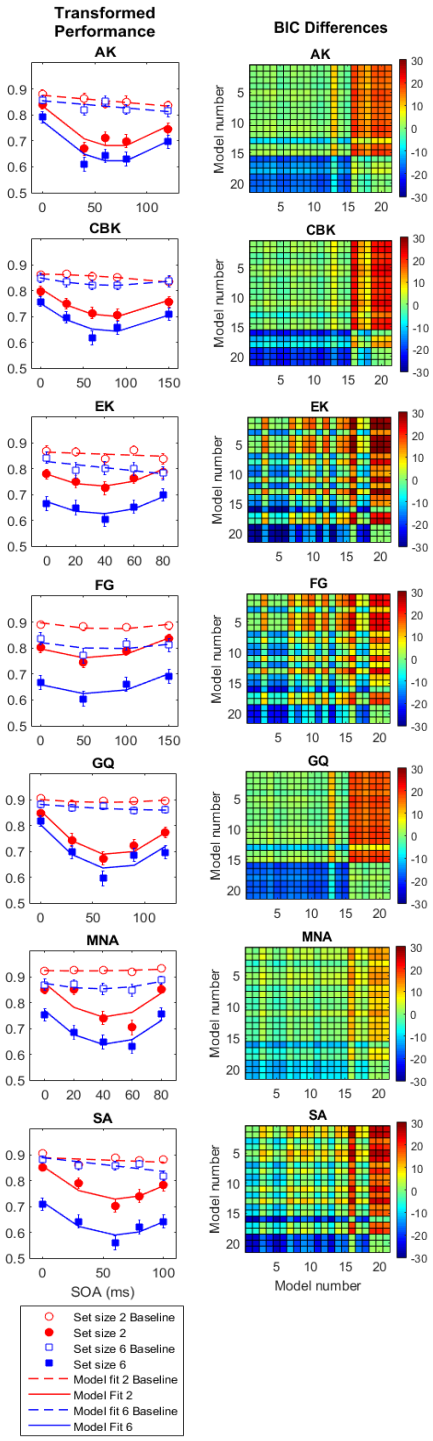

Figure 3-2 The left column shows transformed performance for each observer against SOA. The right column shows pairwise BIC differences[7] between regression models listed in Table 3-II in explaining transformed performances.

[7] A square with coordinates (x,y) on each plot represents BIC difference between model x and y (i.e., BIC_{My} - BIC_{Mx}). The smaller the BIC, the better the model performs, therefore negative values (i.e., cooler colors) represent better model performance. Model M16 was the best model for all observers. Note that adding more terms to model M16 does not improve model performance, which is evident by dark blue bands formed in the lower left quadrant of each plot.

functions appear to be shifted vertically as a function of set-size, mirroring attentional affects found in the transformed performance data. In order to quantify these informal observations, we fitted a series of regression models listed in Table 3-II (see Methods for details).

Pairwise comparison results of all regression models are given in the two rightmost columns of Figure 3-3. For the standard deviation parameters, the model M21 (with the following factors: SOA, SOA^2, set size, SOA x set size, and SOA^2 x set size) outperformed all other regression models (21^{st} rows in in each panel in the fourth column of Figure 3-3) for observers CBK, FG, and SA. For observers AK, EK, and GQ, models M1, M4, and M2 were the best ones, respectively. However, almost all BIC differences were within the range of [-2, 2], suggesting that the differences between the models were not significant and all models performed equally well (or equally poorly). For observer MNA, the model M8 appeared to be the best of all, suggesting significant roles for SOA and set size as well as their first order interaction. In sum, these findings support the aforementioned informal observations that there is no clear or consistent trend across observers in the dependence of the standard deviation parameter on SOA and set-size. In our previous work (Agaoglu et al., 2015), we found that both the standard deviation of the Gaussian and the weight of the uniform distribution in the GU model correlated with the metacontrast function. The correlation of the weight parameter was higher than the correlation of the standard deviation. In the current study, the best regression model for the standard deviation had a main factor of SOA in five out of seven observers, which suggests a significant role for standard deviation of the Gaussian term in explaining

metacontrast masking, consistent with the previous finding. However, this dependence did not show a consistent pattern across observers and hence may be related to individual observer-dependent variations. On the other hand, as we mentioned above and discuss below in more details, the weight parameter appears to reflect a more general property that is common across all observers.

The weight of the uniform component in the GU model showed an inverse U-shaped pattern which was consistent across observers. In three (AK, MNA, and SA) of seven observers, M16 performed best, indicating no interaction between SOA and set size. Interestingly, in the remaining four observers (CBK, EK, FG, and GQ), the best regression model was either M19 or M21, both of which have interaction term(s). The interaction between quadratic SOA term and set size is most apparent in observer CBK (the second row and third column in Figure 3-3). However, for observers EK, FG, and GQ, even though the best regression model is M21, the model M16, which does not have any interaction terms, performs equally well according Jeffrey's scale of interpretation. In fact, regressions of the weight of uniform based on adjusted R^2 metric revealed that M16 is the best regression model for all observers but CBK and EK. Besides, qualitatively, interaction between SOA and set size is not very apparent for these observers. Therefore, we conclude that, although there is some evidence for interactions between masking and attention when the analysis is carried out through the weight of the uniform distribution in the GU model, the evidence for this interaction is neither consistent across observers, nor strong. Hence, in the light of the analysis carried out directly on transformed performance, we conclude that attention and metacontrast masking do not interact. Table

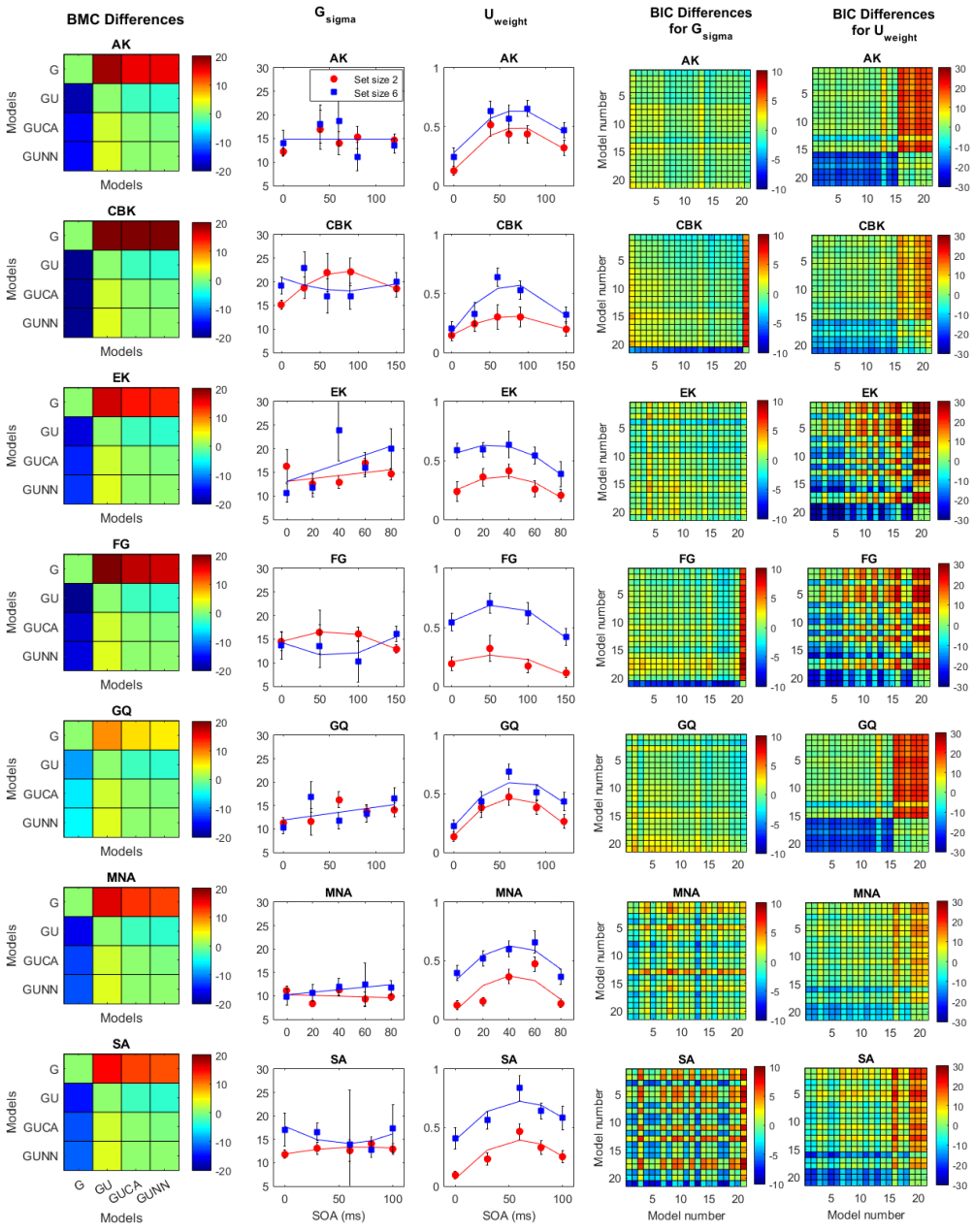

Figure 3-3 The first column from left represents the BMC differences between every combination of model pairs[8]. The second and third columns show the parameters of the winning GU model[9]. The fourth and fifth columns show BIC differences between pairs of regression models.

[8] In order to have the same color notation (i.e., cooler colors mean better model performance and hotter colors mean worse model performance) as in Figure 3-2, we flipped the sign of BMC differences. The GU model outperforms all others for all observers.

[9] The second column shows the standard deviation of the Gaussian in the GU model as a function of SOA, and the third column shows the weight of the Uniform component in the GU model. The red lines represent set size 2 condition whereas the blue lines represent set size 6 condition. Error bars represent standard errors obtained by bootstrapping (see Methods).

3-IV summarizes the best regression models in capturing the change in model parameters as a function SOA and set size, for each observer.

Table 3-IV The winning regression model for each parameter and observer.

Observer	G_{sigma}	U_{weight}
AK	M1	M16
CBK	M21	M21
EK	M4	M19
FG	M21	M21
GQ	M2	M19
MNA	M8	M16
SA	M21	M16

Another way to understand how model parameters and masking functions are related is to compute the correlation between each model parameter and masking functions. Figure 3-4 shows individual correlation coefficients as well as the average across observers. The weight of the Uniform in the GU model strongly correlates with masking functions, and set size does not change the strength of this correlation (one sample t-tests, $p < 0.0001$; Bayes factor $> 7x10^5$, in favor of strong correlation). Interestingly, the standard deviation of the Gaussian in the GU model correlates with masking function in set size 2 condition (one sample t-test: $t_6=-2.50$, $p=0.046$; Bayes factor (correlation/no correlation) = 2.01) but this correlation vanishes in set size 6 condition (one sample t-test: $t_6=-0.67$, $p=0.53$; Bayes factor = 0.42.) We will discuss these findings in the next section.

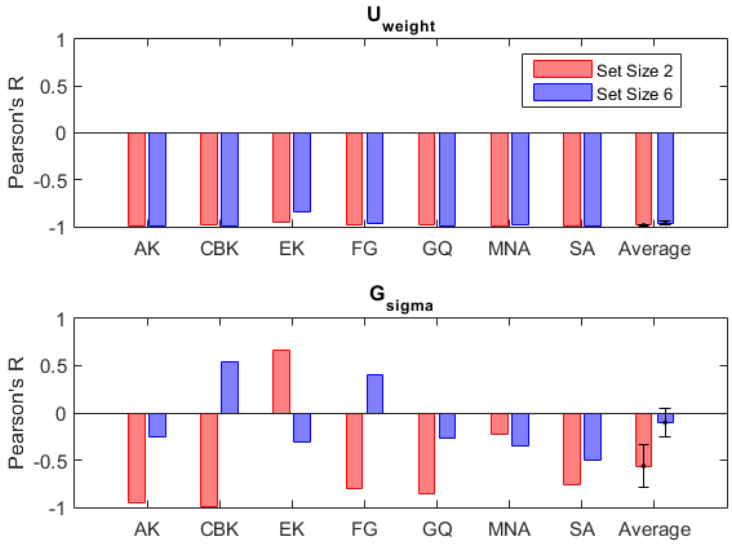

Figure 3-4 The correlation between model parameters and masking functions for each set size condition. The correlation coefficients for individual observers as well as average across observers are shown. The red and blue bars represent set size 2 and 6 conditions, respectively.

3.4. Discussion

The visual system constantly receives an overwhelming amount of information. Due to capacity limitations, it becomes necessary to select and/or enhance relevant information while suppressing irrelevant information for the task at hand. These attentional effects can be quantified experimentally with tasks that require the observer to detect, discriminate, or recognize a given object. In spatial cueing paradigms, attentional resources are directed to specific spatial locations and performance at cued and uncued locations are compared. In visual search paradigms, the "attentional load" is manipulated by means of different number of distractor objects/features (see review Carrasco 2011 for a detailed taxonomy of attentional effects). Visual masking has also been shown to control the quantity and quality of information transfer from sensory memory to short-term memory. An intuitive question is whether these two processes that control the

transfer of information from sensory memory to short-term memory operate independently or interact.

In this study, we asked observers to report the orientation of a target bar randomly selected from a set of bars presented in the display. Since the target bar was indicated by a metacontrast mask or a peripheral post-cue, we assumed that by increasing the set size, observers spread their attention to more locations thereby reducing attentional benefits at individual locations. We found strong evidence against interactions between metacontrast masking and attentional mechanisms. Our results showed that mean absolute response-errors in orientation judgments are independently influenced by masking strength (a function of SOA) and attentional load (a function of set size).

As mentioned in the Introduction section, while some models of masking view attention as an integral component of masking effects, others view it as an independent add-on process. In particular, the object-substitution model of masking, which was derived from the common-onset masking experimental paradigm, posited interactions between masking and attention and provided empirical evidence in support of this prediction. Other studies provided empirical evidence for masking-attention interactions in metacontrast masking (Ramachandran & Cobb, 1995; Shelley-Tremblay & Mack, 1999; Tata, 2002) raising the possibility that these interactions could be an essential component of *all* masking types. However, recent studies, using the common-onset masking paradigm, showed that the interaction between masking and attention was an artifact of ceiling/floor effects and provided evidence against the prediction of the object substitution model (Argyropoulos et al., 2013; Filmer et al., 2014, 2015; Pilling et al., 2014). A goal of our study was to examine whether the interaction between attention and

masking in metacontrast could also be a result of floor/ceiling effects. By avoiding floor/ceiling effects, we showed strong evidence against masking and attention interactions in metacontrast masking. In the light of this finding, we now discuss previous studies that reported interactions between these two processes.

Ramachandran and Cobb (1995) used a row of three disks (central one being the target) and a column of four flanking disks (two above and two below the target disk). They asked observers to give a visibility rating for the target disk on a scale of 0 to 5. They found *stronger* masking when observers attended the column of disks which constituted the mask compared to when they attended the row of disks that included the target. The authors interpreted this finding as an interaction between attention and backward masking. However, it is very likely that the interaction reported by Ramachandran and Cobb was a result of ceiling effect: When observers attended the row of disks containing the target, visibility ratings were high, and for some SOA values, were very close to 5 (the maximum value).

Tata (2002) reported similar findings and interpretations with metacontrast masking. He used elements similar to Landolt Cs and asked observers to report the orientation of the masked one. He varied set-size to control the attentional load and found significant interactions between set-size and masking. However, as in Ramachandran and Cobb's study, performances in Tata's experiments also suffered from ceiling effects: For short and long SOAs (e.g., 0 ms and 240 ms), discrimination performance in all set-size conditions was in the range of 90-95% correct whereas at intermediate SOA values, performance dropped significantly and diverged. The ceiling effect was rather more obvious in this study because with a set-size of one, there was essentially no masking at

all (performance as a function of SOA formed a flat line at about 95% correct), whereas at set-size of eight, there was strong masking with a typical type-B masking function.

Another study that investigated metacontrast masking and attention also showed significant interactions (Shelley-Tremblay & Mack, 1999). In inattentional blindness studies, meaningful stimuli were found to be more resistant to inattentional blindness than neutral stimuli. This was interpreted as meaningful stimuli automatically attracting additional attentional resources compared to neutral stimuli. Following this logic, Shelley-Tremblay & Mack (1999) manipulated attention by using meaningful (happy-face icon, individual name) vs neutral stimuli (inverted face icon, scrambled face icon, neutral words, annulus). They found that targets consisting of happy-faces and one's own name were more resistant to masking than scrambled variants of them (facial features within a happy-face icon or letters in one's name were randomly located) and meaningful stimuli used as masks exerted stronger masking effects than neutral masks. More importantly, their data indicated significant interactions between target/mask manipulations and SOA. The interpretion of these data in favor of interactions is subject to two important caveats: First, baseline performance for each type of stimulus (i.e., without a mask) was not measured; therefore one cannot judge the strength of masking and/or the presence of a ceiling effect. Second, in the experimental design, target or mask type *covaries* with attentional manipulation. This is especially important given that changes in the target or mask type, not only in terms of low-level parameters (e.g., luminance), but also in terms of higher-level organization, are known to affect metacontrast masking functions (e.g., Dombrowe, Hermens, Francis, & Herzog, 2009; Sayim, Manassi, & Herzog, 2014; Williams & Weisstein, 1981). For example, in two

studies (A. Williams & Weisstein, 1978; M. C. Williams & Weisstein, 1981), target and mask configurations are manipulated in terms of their three-dimensional appearance and connectedness. Both of these factors affected metaconstrast functions; connectedness influencing mainly the *strength* of masking whereas depth influencing mainly the *timing* of masking. Similar types of influences would be expected in the case of Shelley-Tremblay's & Mack 's stimuli: Given the cognitive significance of happy faces and one's own name, it is likely that they are processed faster than neutral stimuli, suggesting shifts in the timing of metacontrast, hence interaction effects. In summary, because the target or the mask type covaried with the attentional manipulation, it is not clear whether the interaction found in Shelley-Tremblay & Mack (1999) is due to target and mask types based on figural, Gestalt, or "object superiority" effects, or to attention itself.

3.4.1. Effects of Attention and Masking on Signal and Noise

Although attentional effects are very well established with various visual tasks, there is no consensus about its mechanistic bases. Based on psychophysical, neurophysiological, and neuroimaging data, many computational models of attention have been proposed. Proposals include signal enhancement, external noise reduction, distractor exclusion, change in decision criteria and/or spatial uncertainty, normalization of pre-attentive activity by attention/suppression fields, increase in information transfer to VSTM, accelerating information processing, sharpening of tuning curves, modulating contrast and/or response gains, and many more (e.g., Carrasco & McElree, 2001; Carrasco, Penpeci-Talgar, & Eckstein, 2000; Carrasco, 2011; Desimone & Duncan, 1995; Dosher & Lu, 2000a, 2000b; Eckstein, 1998; Herrmann, Montaser-Kouhsari, Carrasco, & Heeger, 2010; Lee, Itti, Koch, & Braun, 1999; Lu & Dosher, 1998; John Palmer, 1994;

Pestilli, Ling, & Carrasco, 2009; Reynolds & Heeger, 2009; Smith, Ellis, Sewell, & Wolfgang, 2010; Smith, Lee, Wolfgang, & Ratcliff, 2009; Smith & Ratcliff, 2009). These processes are not mutually exclusive and can work in parallel with different contributions in different stimulus/task conditions. For instance, in precuing of location, the effects of cue-validity can be explained primarily by external noise reduction when there is high amount of noise in the stimuli whereas signal enhancement accounts for attentional effects in low external noise conditions (Dosher & Lu, 2000a; Lu & Dosher, 1998). Modulating contrast and response gains have been associated with endogenous (i.e., central cueing), and exogenous (i.e., peripheral cueing) attention, respectively (Herrmann et al., 2010; Pestilli et al., 2009). What do our results imply in terms of signal and noise modulation by attention and masking? Our data suggest that masking reduces the target signal-to-noise ratio (SNR) whereas decreasing attentional load increases it and their effects simply add up. A simple interpretation of our results is that the metacontrast mask reduces the strength of the target signal thereby reducing SNR whereas attention enhances signal strength, given that our target is presented under low noise conditions. Given the lack of interactions between metacontrast and attention, these signal enhancement and reduction modulations by masking and attention take place as independent additive effects.

3.4.2. Implications for Models of Attention

Lu and Dosher developed a theoretical and experimental framework to investigate potential mechanisms of attention (Lu & Dosher, 1998). According to this framework, three distinct mechanisms of attention can be differentiated experimentally by adding varying levels of noise to the visual stimuli. The Perceptual Template Model (PTM)

consists of four stages and incorporates both additive and multiplicative noise sources. The first stage is a "perceptual template", modeled as a filter tuned to the signal. This stage filters out some of the external noise that accompanies the desired signal. In the second stage, the output of the first stage is rectified and fed into a multiplicative Gaussian noise source with zero mean and a standard deviation proportional to the signal strength (i.e., its total energy). In the third stage, an independent Gaussian noise with zero mean and a constant standard deviation is added. The last stage is a standard signal detection (i.e., decision) process that is appropriate to the task and the stimuli.

PTM can differentiate three distinct attention mechanisms each of which leads to a signature behavioral improvement in perceptual tasks. These mechanisms are (i) stimulus enhancement, (ii) external noise exclusion, and (iii) multiplicative noise reduction. There are both physiological and behavioral evidence in support of these mechanisms. For instance, at the neurophysiological level, attention has been shown to increase cellular response sensitivity (Reynolds & Chelazzi, 2004; Reynolds et al., 2000), to sharpen tuning curves of orientation and spatial frequency selective cells (Haenny et al., 1988), and to shrink neuronal receptive fields thereby excluding unwanted information through intra- or inter-layer interactions (Desimone & Duncan, 1995). At the behavioral level, attention has been associated with reduction in decision uncertainty (Palmer et al., 1993), enhancement of the attended stimuli (Lu & Dosher, 1998; Lu et al., 2000; Posner et al., 1978), exclusion of external noise or distractors (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 2000; Lu et al., 2002; Shiu & Pashler, 1994), and modulation of contrast-gain (Lee et al., 1999).

There are two broad categories of spatial cueing, namely central and peripheral cueing. Central cues are generally presented at the locus of fixation and signal the location of the target stimulus in a way that requires interpretation. For example, when an arrow is used, the observer has to interpret the direction of the arrow to infer the cued location. Central cueing activates voluntary, or endogenous, attention mechanisms. Peripheral cues are generally presented at or close to the spatial location of the stimulus and hence they indicate the location of the stimulus directly in spatial representations without necessitating interpretive processes. These cues activate the reflexive, or exogenous, attention mechanisms. Lu and Dosher (2000) found that endogenous attention works by external noise exclusion whereas exogenous attention invokes both external noise exclusion and signal enhancement mechanisms.

We will consider whether PTM can explain our findings. In our experiment, we have manipulated set-size to control attention. In this case, the main type of attention in effect would be endogenous attention, since observers spread their attention voluntarily to more locations as set-size increases. PTM predicts that external noise exclusion is the mechanism underlying endogenous attention effects. Under the external noise exclusion scenario, PTM predicts large attentional effects when external noise is large. If the mask's effect is to add noise to the stimulus, then more noise should have been added when masking is strong (e.g., Lu, Jeon, & Dosher, 2004). Accordingly, the effect of attention should be strong when masking is strong and weak when masking is weak, hence there should be interactions between attention and masking. This does not agree with our results.

Several studies reported that cuing improves sensitivity in simple detection tasks when stimuli are presented with masks but not when stimuli are presented in the absence of masks (e.g., Lu & Dosher, 1998, 2000; Lu et al., 2002; Smith & Wolfgang, 2004, 2007). Smith and colleagues developed the integrated system model (ISM) to explain these findings (Smith & Wolfgang, 2004 – early version, no explicit VSTM layer; Smith & Ratcliff, 2009 – VSTM stage is added; Smith et al., 2010 – final version). The main assumption of the model is that attention affects the rate of information transfer from sensory memory to VSTM (Carrasco & McElree, 2001) . Crucially, ISM incorporates interacting masking and attention mechanisms and predicts larger attentional benefits when a stimulus is masked compared to when it is unmasked. Likewise, the stronger the masking is, the larger the attentional effects will be. Hence, both the aforementioned empirical findings and the predictions of ISM appear to be at odds with our findings: Our baseline data, which correspond to no mask conditions, show clear effects of attention and we found no interactions between attention and masking. However, it is important to point out that the experimental paradigms leading to different results are fundamentally different: Lack of attentional effects for unmasked stimuli were found for *simple detection tasks* (or equivalently for easy discrimination tasks, such as horizontal vs vertical) that are mainly limited by contrast, rather than by the similarity of stimulus alternatives. This is clearly not the case in our study, wherein observers are required to report as accurately as possible the orientation of the target line. Hence, we found the classical set-size effect in our no-mask baseline conditions, in agreement with other studies (e.g., John Palmer, 1994). It is well known that the magnocellular pathway and its associated transient mechanisms have very different contrast responses compared to

parvocellular pathway and its associated sustained mechanisms (Croner & Kaplan, 1995; Kaplan & Shapley, 1986). Simple detection and easy discrimination tasks can be carried out by both transient and sustained mechanisms, whereas difficult fine-discrimination tasks are likely to necessitate sustained mechanisms. Hence, both task difficulty and the contrast level are expected to influence the mechanistic criterion *contents,* i.e., which mechanisms, sustained or transient, will contribute to performance. Given that attention is also known to influence transient and sustained mechanisms in different ways, the interaction effects that emerge from data may be due to changes in criterion contents. In fact, this is a major challenge for any study, including ours, seeking to analyze interactions of masking with other processes. Masking is not a unitary phenomenon and different criterion contents can lead to drastically different masking functions. In order to mitigate this issue, in this study we sought to analyze interactions based on a complete type-B metacontrast function comparing identical masking conditions (i.e., identical SOAs) while modulating attention via set-size.

To summarize, masking and attention are both involved in information processing and transfer at multiple stages of visual processing. Determining their relationships can help us reach a richer and more integrated understanding of visual information processing. Previous studies showed significant interactions between different types of masking and attention. However, in most of these studies, findings suffered from methodological artifacts and/or could be interpreted by alternative accounts (rather than artefactual mask-attention interaction). Here, we investigated the relationship between metacontrast masking and attention based on two performance measures: (i) mean absolute response-errors (empirical), (ii) distribution of signed response-errors

(modeling). We found strong evidence against interactions between attention and metacontrast masking for both performance measures. As mentioned above, neither masking nor attention is a unitary phenomenon, and hence additional studies are needed to establish firmly the relations between *types* of masking and attention.

Chapter 4. Temporal Dynamics of the Effect of Endogenous and Exogenous Attention on Visual Masking[10]

4.1. Introduction

The visual system is flooded with an enormous amount of information under normal viewing conditions. Only a subset of this information can be selected for further processing. Attentional mechanisms are responsible for enhancing the processing of the selected information (items, objects, etc.) and suppressing (or filtering out) the rest by allocating the available processing resources accordingly. The selection and filtering functions of visual attention have been investigated extensively and are well-documented (e.g., Chen et al., 2008; Gazzaley & Nobre, 2012; Palmer, 1990; Polk et al., 2008). In short, attention modulates the information transfer from sensory memory to VSTM and it has a significant role in the maintenance of information in VSTM (Ogmen et al., 2013; Reynolds & Chelazzi, 2004; Sreenivasan & Jha, 2007; Tombu et al., 2011). Moreover, two distinct types of attentional orienting have been identified (Cheal & Lyon, 1991; Egeth & Yantis, 1997; Jonides, 1981; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Posner, 1980; Weichselgartner & Sperling, 1987): Exogenous attention has often been described as controlled by the stimulus and, therefore, referred to as a reflexive mechanism. When we hear a loud bang or see a flash of light on a dark road, the visual system automatically deploys most of its resources for processing this information. Hence, exogenous attention has a significant role in survival. Endogenous attention, on

[10] All of the findings reported in this chapter will be submitted for publication as a journal paper. Therefore, the style and the content follow the journal guidelines.

the other hand, is a voluntary, rather than reflexive, allocation of resources to a predetermined region in the space, to a particular feature, or to an object.

Peripheral cues (presented at or around the target stimulus) have been claimed to activate exogenous attention whereas central cues (generally presented at or near fovea) are said to trigger endogenous attention mechanisms. Peripheral cues directly specify the target location whereas central cues are conceptual in the sense that they need to be cognitively processed and interpreted to determine where and how to deploy attentional resources. Due to these differences, exogenous attention reaches its maximum effectiveness at shorter cue-target onset asynchronies (CTOA) (100-120 ms depending on the task and the stimuli) compared to endogenous attention, which may require about 300 ms to reach its maximum effectiveness (e.g., Ling & Carrasco, 2006; Remington, Johnston, & Yantis, 1992). Moreover, exogenous attention effects decrease and disappear completely after 300-400 ms whereas endogenous attention benefits show a monotonically increasing trend as a function of CTOA and can be maintained as long as are needed for the task (Cheal & Lyon, 1991; Kröse & Julesz, 1989; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; see review: Carrasco, 2011). Due to these differences in time courses, exogenous and endogenous attention have been also called as transient and sustained attention respectively (Egeth & Yantis, 1997; Jonides, 1981; Nakayama & Mackeben, 1989; Wright & Ward, 2008).

Endogenous attention and exogenous attention can result in similar perceptual changes (e.g., Hikosaka, Miyauchi, & Shimojo, 1993). Both types of attention have been shown to increase spatial resolution (Carrasco & Barbot, 2015; Yeshurun & Carrasco, 1998) and reduce temporal resolution (Yeshurun & Levy, 2003) at the attended location,

and affect performance in various behavioral tasks such as visual search (e.g., Carrasco & McElree, 2001), crowding (Montaser-Kouhsari & Carrasco, 2009; Yeshurun & Rashal, 2010), and acuity (Montagna, Pestilli, & Carrasco, 2009). Suzuki and Cavanagh (1997) showed that both types of attention distort the representation of position at the attended location. On the other hand, they can also result in distinct perceptual effects. The effect of exogenous attention on conjunction search (based on the conjunction of multiple features) is larger than on simple search (based on a single feature) whereas endogenous attention yields equivalent improvements (Briand & Klein, 1987). Dosher and Lu, in a series of studies, showed that endogenous attention operates only under high-noise conditions whereas exogenous attention benefits can be found under both low-noise and high-noise conditions (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000). Ling and Carrasco (2006), however, showed that both types of attention increase contrast sensitivity in both high- and low-noise conditions. Moreover, modulating contrast and response gains of neuronal responses have been associated with endogenous and exogenous attention, respectively (Herrmann et al., 2010; Pestilli et al., 2009). Due to varying differences and similarities in temporal dynamics and perceptual changes they produce, there seems to be no consensus about the underlying neural mechanisms of these two types of attention. The view that the neural networks underlying endogenous and exogenous attention overlap to some extent but are independent, has been supported by many studies (Carrasco, 2011) with one exception. Peelen et al. (2004) used both central and peripheral cues in a functional neuroimaging study, and reported that the same large-scale neural network (a fronto-parietal network consisting of premotor cortex, posterior parietal cortex, medial frontal cortex and right inferior frontal cortex) mediates

both types of attention. Nevertheless, whether masking has the same relationship with these two types of attention remains to be established empirically.

Attentional allocation of resources can also be controlled by changing set size rather than by a spatial cue, as we have shown in Chapter 3. However, there are two limitations of controlling attentional allocation this way. First, it does not allow us to investigate the temporal dynamics of attentional benefits. Second, since observers had to attend to the entire display in the beginning of each trial and the target was indicated by the onset of a mask, the task employs both endogenous and exogenous attention. Due to different temporal dynamics of the two types of attention, one cannot tease apart their contributions on performance. Moreover, differential contributions from different attention mechanisms might have shadowed a potential interaction between metacontrast masking and attention in Chapter 3. Here, we investigated the relationship between metacontrast masking and these two different types of attention separately by presenting either central or peripheral cues in different blocks. By adjusting the stimulus parameters, we again made sure that the ceiling/floor artifacts are avoided for all observers. Finally, we used the same statistical modeling approach to determine whether endogenous and exogenous attention give rise to similar changes in the distribution of response errors.

4.2. Methods

4.2.1. Participants

Six observers (three male, three female; ages range from 24 to 32) took part in this study and four of them were naïve as to the purpose of the study. All participants had normal or corrected-to-normal vision, and gave written informed consent before the

experiments. All experiments were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). We followed a protocol approved by the University of Houston Committee for the Protection of Human Subjects.

4.2.2. Apparatus

Stimuli were created using the ViSaGe and VSG2/5 cards manufactured by Cambridge Research Systems. A 22-in. CRT monitor with a refresh rate of 100 Hz and a display resolution of 800 by 600 pixels was used to present the visual stimuli. Observers sat at a distance of 1 m from the display. In order to restrict movements of the observer, a head/chin rest was used. Responses from observers after each trial were collected via a joystick.

4.2.3. Stimuli and Procedures

In order to investigate the interactions between metacontrast masking and the two types of attention described above, we used either a central cue (endogenous attention experiment) or a peripheral cue (exogenous attention experiment) in separate blocks. The task of the observers was to report the orientation of a target bar indicated by the pre-cue at the beginning of each trial. The stimulus sequences for both cueing types are given in Figure 4-1. Each trial started with a fixation point on an otherwise blank gray (60 cd/m^2) screen. After a random delay (500-1000 ms), a black (10 cd/m^2) pre-cue (an arrow at the center in the endogenous attention blocks, or a 0.3 deg square at 3.0 deg eccentricity in the exogenous attention blocks) was shown for 50 ms, indicating the target location. After a variable CTOA, an array of six (endogenous) or four (exogenous) oriented bars was presented for 10 ms around an imaginary circle centered on the fixation point so that

all bars had the same retinal eccentricity. In the endogenous attention blocks, the eccentricity of each bar was 6 deg whereas in the exogenous attention blocks, the eccentricity of each bar was 5 deg. In contrast to the experiments in Chapter 3, all oriented bars were followed by a metacontrast mask (a ring with inner and outer diameters of 1.1 deg and 1.4 deg, respectively) after a variable delay. This variable delay between the onsets of the bar and the mask arrays will be referred to as stimulus onset asynchrony (SOA) in the following text. Within the same block, the mask array was not presented in some trials, and the performance in these trials served as the baseline performance level for each cueing condition. Once the stimulus sequence was presented, observers reported the orientation of the bar indicated by the central cue in the endogenous attention condition whereas they reported the orientation of the bar indicated by another cue (i.e., the post-cue) which appeared at the end of each trial in the exogenous attention condition, by adjusting the orientation of a central bar via a joystick. In other words, in the endogenous attention experiment, the central cue had 100% validity. In the exogenous attention experiment, however, the peripheral cue was not informative of the target location (25% validity). This ensured that any potential contribution from endogenous attention mechanisms is minimized. There was no time limitation on the response, and observers initiated the next trial by another button press. In the endogenous attention experiment, three CTOA values (0, 200, and 500 ms) were used whereas in the exogenous attention experiment, only two CTOAs were used; 0 ms and a CTOA between 80 ms and 120 ms (specific values for each subject were as follows: 120 ms for ATB, 80 ms for EB, 100 ms for FG, 80 ms for GQ, 100 ms for MNA, and 80 ms for SA), where the effect of exogenous attention is largest, as

determined by the pilot studies. In both cueing conditions, five SOA values were used for each observer.

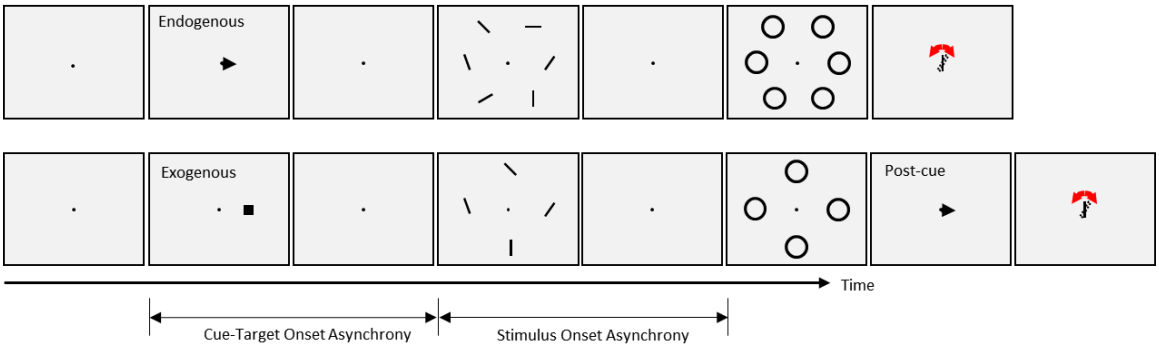

Figure 4-1 The stimulus sequences for both the endogenous (top) and exogenous (bottom) attention conditions.

In both experiments, each block started and ended with 10 consecutive baseline trials. Moreover, different SOAs were interleaved in the remaining trials within a block. In each block, the same CTOA was used. In other words, cue timing was blocked whereas target and mask timing was randomized. Each combination of CTOA and SOA values as well as the baseline conditions were run 100 times. In total, each observer completed 1800 trials ([5 SOA + 1 baseline] x 3 CTOA = 18 conditions) in the endogenous attention experiment, and 1200 valid trials ([5 SOA + 1 baseline] x 2 CTOA = 12 conditions) out of roughly 4800 trials (i.e., 25% validity) in the exogenous attention experiment.

As in the previous chapters, in order to obtain masking functions, response errors were transformed to a probability-like measure via Equation (4-1) (Ogmen et al., 2013).

$$Transformed\ Performance = 1 - \frac{|Error\ Angle|}{90} \quad . \qquad (4\text{-}1)$$

Moreover, as for the set-size experiment in Chapter 3, the target and mask luminances were adjusted for each observer to avoid floor/ceiling artifacts. In short, we used two criteria to select the target-mask luminance pairs: C1) the luminance pairs which resulted in significant differences between masked and unmasked (baseline) performance, and C2) the luminance pairs which resulted in significant differences between masked performance and the chance level, for all SOAs. Moreover, the SOAs were also selected for each observer separately so that typical type-B masking functions can be captured. Table 4-I lists the target and mask luminances, as well as the results of t-tests used to check whether both criteria listed above were met, for all observers.

4.2.4. Statistical Analyses and Modeling

The analyses of transformed performance and the statistical modeling of the distribution of signed response errors were identical to those in Chapter 3. In the exogenous attention condition, only the trials where the peripheral cue correctly indicated the target location (i.e., valid trials) were included in the analyses.

4.2.5. Predictions

As we mentioned in the Introduction section, endogenous and exogenous orienting have been known to have different time courses in enhancing target processing (see review: Ward, 2008). Figure 4-2A illustrates the time courses and the predicted outcomes for the experiments presented here. When either type of cue is shown simultaneously with the target item (i.e., CTOA = 0 ms), both cues are ineffective (Cheal & Lyon, 1991; Kröse & Julesz, 1989; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989). However, as the time separation between the cue and the target is increased, the

Table 4-I The target, mask, and cue luminance values in cd/m2 (and corresponding Weber contrasts) for each observer. The background luminance was set to 60 cd/m2 for all observers. The results of t-tests used to determine whether criteria C1 and C2[11] are met are also listed.

Endogenous

	Luminance (Contrast)			Statistical Criteria	
Observer	**Target**	**Mask**	**Cue**	**C1 (ceiling)**	**C2 (floor)**
ATB	*43 (-0.28)*	*15 (-0.75)*	*10 (-0.83)*	t(150.7) = -2.18; p = 0.016	t(99) = 4.11; p < 0.001
EB	*12.5 (-0.79)*	*30 (-0.5)*	*10 (-0.83)*	t(164.8) = -2.22; p = 0.014	t(99) = 6.15; p < 0.001
FG	*46 (-0.23)*	*18 (-0.7)*	*10 (-0.83)*	t(145.5) = -2.73; p = 0.004	t(99) = 8.27; p < 0.001
GQ	*46 (-0.23)*	*0 (-1)*	*10 (-0.83)*	t(141.6) = -2.53; p = 0.006	t(99) = 3.92; p < 0.001
MNA	*42 (-0.3)*	*20 (-0.67)*	*10 (-0.83)*	t(142.6) = -2.19; p = 0.015	t(99) = 5.86; p < 0.001
SA	*47 (-0.22)*	*18 (-0.7)*	*10 (-0.83)*	t(138.6) = -2.94; p = 0.002	t(99) = 7.72; p < 0.001

Exogenous

	Luminance (Contrast)			Statistical Criteria	
Observer	**Target**	**Mask**	**Cue**	**C1 (ceiling)**	**C2 (floor)**
ATB	*44 (-0.27)*	*10 (-0.83)*	*30 (-0.5)*	t(167.7) = -2.23; p = 0.013	t(99) = 6.26; p < 0.001
EB	*40.5 (-0.32)*	*12 (-0.8)*	*30 (-0.5)*	t(181.5) = -1.87; p = 0.031	t(99) = 5.24; p < 0.001
FG	*46 (-0.23)*	*18 (-0.7)*	*30 (-0.5)*	t(180.6) = -2.5; p = 0.007	t(99) = 6.26; p < 0.001
GQ	*46.5 (-0.22)*	*6 (-0.9)*	*30 (-0.5)*	t(125.3) = -2.9; p = 0.002	t(69) = 4.34; p < 0.001
MNA	*43.5 (-0.28)*	*30 (-0.5)*	*30 (-0.5)*	t(137) = -2.34; p = 0.01	t(99) = 6.68; p < 0.001
SA	*48 (-0.2)*	*30 (-0.5)*	*30 (-0.5)*	t(134.6) = -3.81; p < 0.001	t(99) = 3.73; p < 0.001

facilitatory effect of exogenous attention increases first, peaking around 100-120 ms, and then decreases back to no facilitation at long CTOAs (Cheal & Lyon, 1991; Kröse & Julesz, 1989; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989). For endogenous attention, the facilitatory effect increases monotonically and reaches a plateau after a certain CTOA (Cheal & Lyon, 1991; Kröse & Julesz, 1989; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989). Here, we investigated whether different types of attentional orienting interact with metacontrast masking. If there is no interaction, then

[11] Note that we used two-sample t-tests with unequal variances for testing for C1, and one-sample t-tests against chance level (0.5) for testing for C2.

masking functions (i.e., transformed performance as a function of SOA) should simply shift vertically up or down depending on CTOA. More specifically, masking functions should shift upward with increasing CTOA for the case of endogenous cueing whereas it should shift up first, and then shift down to its no facilitation levels for exogenous cueing (see Figure 4-2B). However, since we used only two CTOAs in the exogenous attention condition, our data can only show an upward vertical shift from zero CTOA to 100 ms CTOA. On the other hand, any other pattern of results such as a change in maximum deviation in masking strength as a function of SOA with CTOA (Figure 4-2C), or a shift of the dip of the masking functions with CTOA (Figure 4-2D), or any combination of these two changes would indicate an interaction between attention and masking.

4.3. Results

Figure 4-3 shows the experimental results for both cueing types for all observers. The vertical axes represent the transformed performance while the horizontal axes represent SOA between the target and mask arrays. The dotted lines represent baseline conditions where the mask array was not presented. The markers and dashed lines represent empirical data whereas the solid lines indicate the best fitting regression models. Different colors represent different CTOAs. The baseline data were collected to ensure that the masking data did not have any ceiling effect (criterion C1, see Methods). For each observer, we performed a two-sample t-test between the baseline and masking conditions at an SOA where masking is the weakest, i.e., the transformed performance is the highest. Moreover, we did a one-sample t-test between the chance level (0.5 transformed performance) and the masking conditions at an SOA and CTOA pair where the transformed performance is the lowest (typically, zero CTOA and an intermediate

SOA), to ensure that floor effects are also avoided (criterion C2, see Methods). Table 4-I summarizes the results of all t-tests as well as the target and mask luminances which allowed us to avoid ceiling and floor artifacts for each observer. In short, both criteria were met for all observers, and our masking data are free from ceiling and floor artifacts.

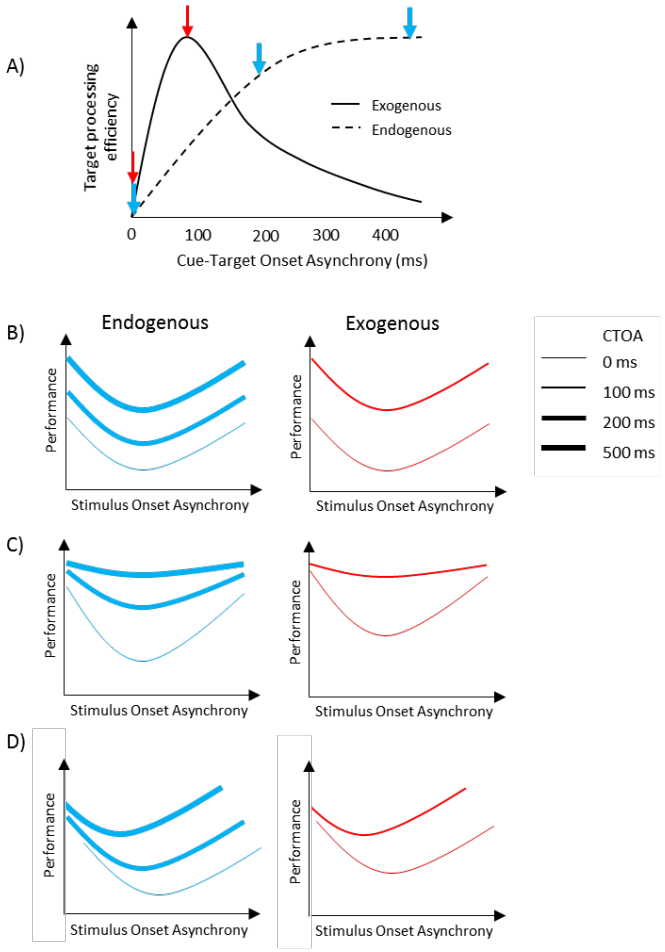

Figure 4-2 (A) The time courses of effects of exogenous (solid line) and endogenous (dashed line) cueing (Ward, 2008). The blue and red arrows indicate endogenous and exogenous cues, respectively. (B) The predicted outcomes assuming no interaction between attention and masking.

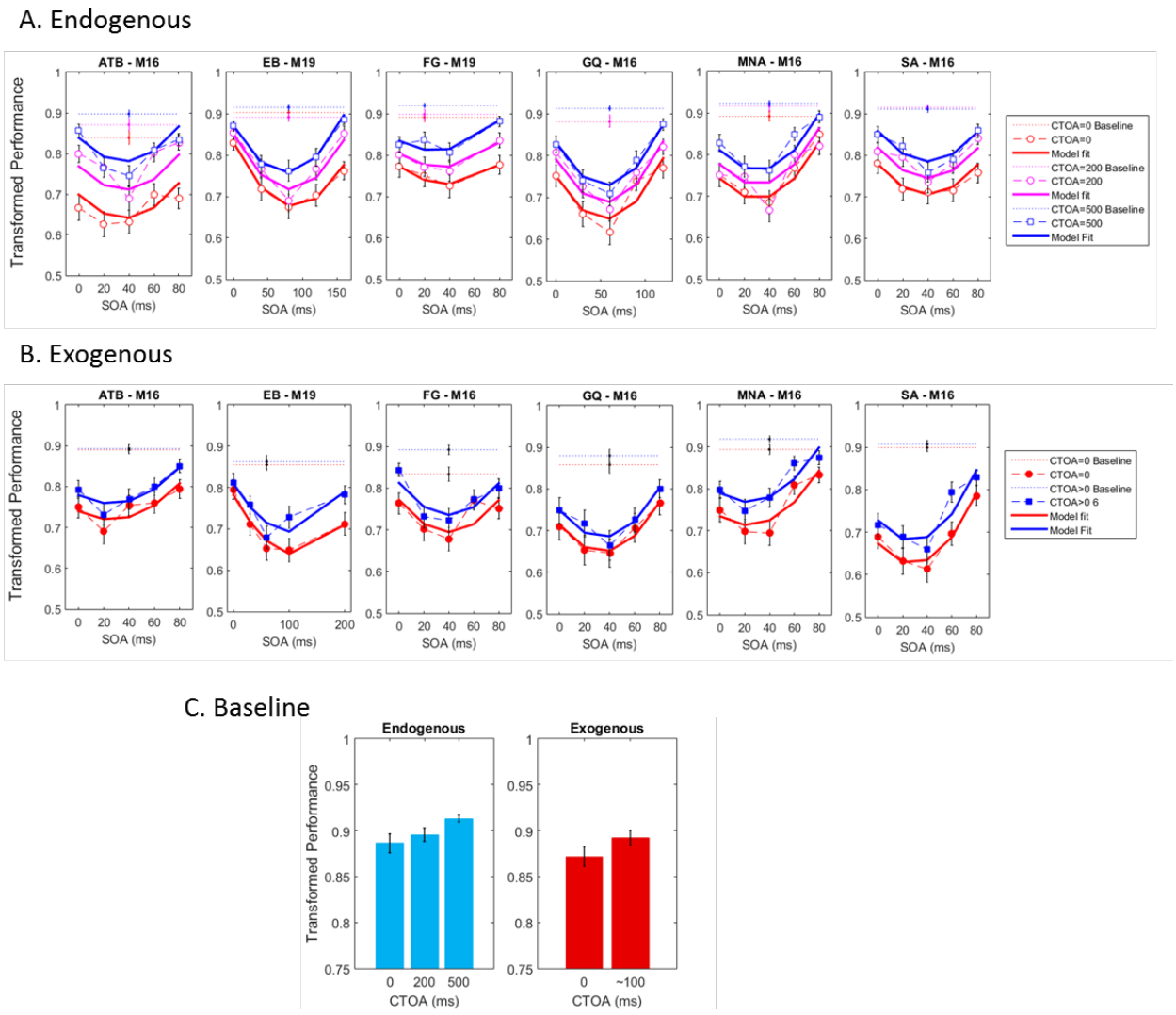

Figure 4-3 The transformed performance in the (A) endogenous and (B) exogenous attention conditions for all CTOAs and SOAs[12]. (C) The baseline performance (averaged across observers) as a function of CTOA in both conditions. Error bars represent ±SEM across observers (n=6).

For both types of attention and for all observers, except EB, masking functions seem to be shifted vertically with changes in CTOA, consistent with the predictions of no

[12]The horizontal axes represent SOA and the vertical axes represent transformed performance (see Methods). Different colors represent different CTOA conditions. The dotted horizontal lines indicate baseline (i.e., without masks) performance. The markers and the dashed lines represent empirical data whereas the solid lines show the regression model which fits the data best. Each panel shows data from a single observer. The initials of each observer and the best regression model (see Table 3-II) are given on top of each panel. Error bars represent ±SEM across trials (n=100). Note that only the validly cued trials are included both conditions, which correspond to 100% and 25% of the trials in the endogenous and exogenous attention conditions, respectively.

interaction between masking and attention. To confirm this qualitative assessment, we fitted a series of polynomial regression models (see Table 3-II in Chapter 3) to individual data to quantify the effects of SOA (masking), CTOA (attention), and their various interactions. The best model was selected based on the BIC metric (the lower the BIC, the better the model) which pits model likelihoods against each other after taking into account the number of parameters. The pairwise BIC differences are given Figure 4-4. Greenish colors represent equivalent model performance whereas blue and red colors represent better and worse model performance, respectively. Figure 4-3 also shows the best model fits (solid lines). For four out of six observers, the best model was M16 in the endogenous attention condition. This model has a linear SOA and CTOA terms as well as a quadratic SOA term but no interaction term. For observers EB and FG, the best model was M19, which has an additional interaction term. In the exogenous attention condition, for all observers except EB, M16 was again the best regression model. For EB, M19 again performed best. Note that although the best regression model was M19 for EB in both attention conditions, and for FG in the endogenous attention condition, the BIC differences between M16 and M19 were within ±2 indicating that these two models actually performed equally well. In sum, our results suggest that metacontrast masking and attention do not interact regardless of the type of cueing used.

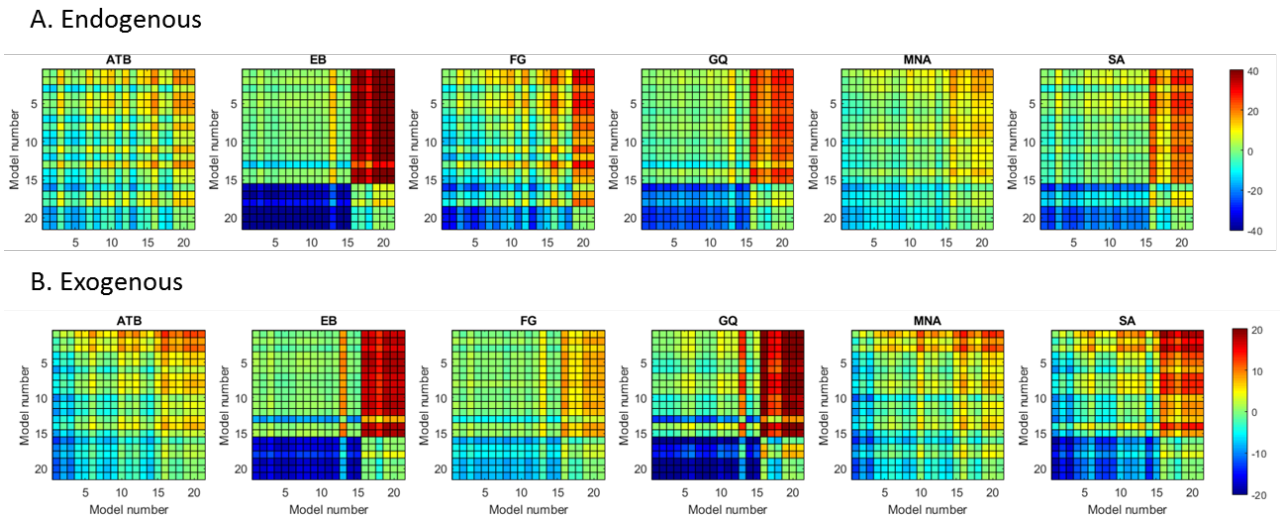

Figure 4-4 The BIC differences between each pair of the regression models listed in Table 3-II in Chapter 3. (A) Endogenous attention condition. (B) Exogenous attention condition.

Perceptual improvements as a result of spatial pre-cueing have been reported to be contingent upon presence of masks (e.g., Lu & Dosher, 1998, 2000; Lu et al., 2002; Smith & Wolfgang, 2004, 2007; Smith, 2000). In order to test whether the effect of cueing is limited to the cases where masks were presented in our experiments, we analyzed the transformed performance in the baseline conditions (see Figure 4-3C). Although we did not control for ceiling and floor effects in the baseline conditions, we found a significant improvement in transformed performance with increasing CTOA in the endogenous attention condition. A one-way repeated measures ANOVA yielded a significant main effect of CTOA ($F_{2,10}$=8.060; p=0.008; η_p^2=0.617). Although there was an increasing trend in performance with CTOA, a paired t-test between performance at zero CTOA and ~100ms CTOA in the exogenous attention condition did not reach significance (t(5)=2.451; p=0.058).

4.3.1. Modeling

We examined the distribution of signed response-errors by using the BMC technique. The same models that were used in Chapter 3 were used. In short, the G model has only a Gaussian term, the GU model is a weighted sum of a Gaussian and a Uniform distribution. The GUCA and GUNN models have an additional Gaussian term, which represents "misbinding" behavior (i.e., reporting the orientation of a non-target object). In the GUCA model, misbinding is caused by the non-target bar which has the closest angle to the target's orientation. In the GUNN model, misbinding is modeled as stemming from the nearest neighbors of the target bar. The rationale for using these models is given in Chapter 3. Among these models, the GU model was the winning model for all observers in both types of attention manipulations. Averaged across observes, the BMC of the GU model in the endogenous attention condition was larger by 13.1, 1.9, and 3.4 than that of the G, GUCA, and GUNN models, respectively. These differences correspond to 5.0E+5-to-1, 6.7-to-1, and 30.0-to-1 odds, in favor of the GU model, and suggest a "decisive evidence" favoring the GU model (Jeffreys, 1998). Similarly, in the exogenous attention condition, the BMC of the GU model was 14.3, 2.1, and 3.2 larger than that of the G, GUCA, and GUNN models, respectively. These BMC differences correspond to 1.6E+6-to-1, 8.2-to-1, and 24.5-to-1 odds, all favoring the GU model. Next, we analyzed the model parameters of the GU model to determine whether any interaction between metacontrast masking and attention exists. Since the Gaussian and the Uniform components in the GU model represent different processes (stimulus encoding and guessing), examination of model parameters have the potential to tease apart different relationships between these processes and attention.

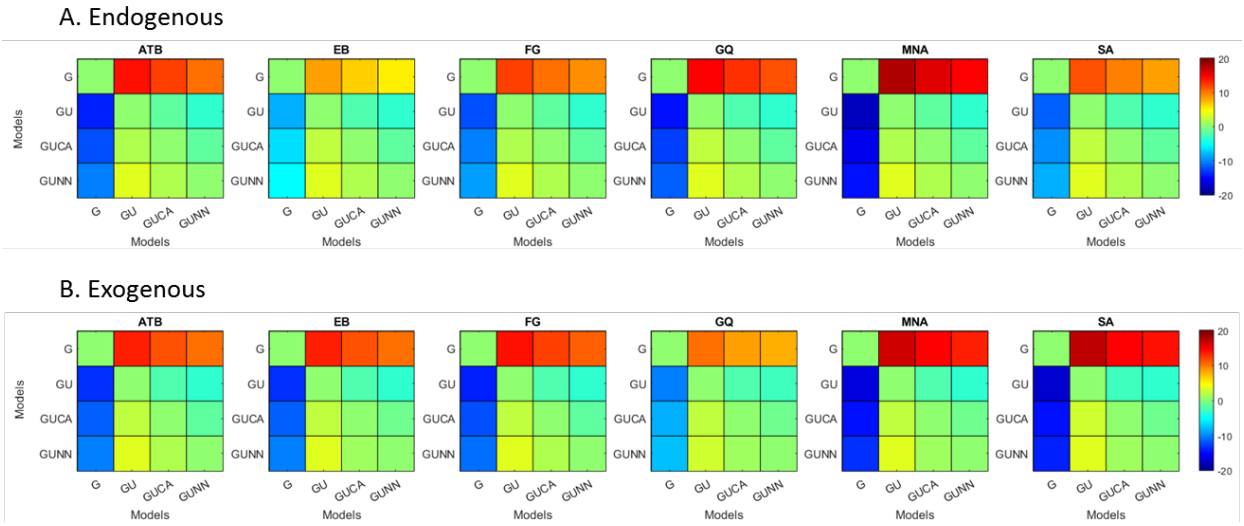

Figure 4-5 Pairwise BMC differences between the statistical models tested. A square with coordinates (x,y) on each plot represents the BMC difference[13] between model y and x.

Figure 4-6 shows the model parameters for the winning GU model as a function of SOA and CTOA in both the endogenous (Figure 4-6A) and exogenous (Figure 4-6B) attention conditions. There is no discernable pattern of changes in the standard deviation of the Gaussian term. The weight of the Uniform component, however, depicts an entirely different picture. First, it tightly follows the (inverted) shape of masking functions, indicating that metacontrast masking exerts its effect by increasing the weight of the Uniform component (i.e., guessing). Second, the effect of pre-cueing at different temporal distances to the target array (i.e., CTOAs) is also reflected in the weight parameter as an overall increase/decrease at all SOAs. At zero CTOA, where spatial pre-cueing virtually has no effect on performance, the weight parameter is largest for all SOAs and observers in both types of attention. As CTOA increases, more attentional

[13] In order to have the same color notation (i.e., cooler colors mean better model performance and hotter colors mean worse model performance) as in Figure 4-4, we flipped the sign of the BMC differences. For both types of attention and for all observers, the GU model performs best in explaining the distribution of signed response errors, as indicated by the darkest blue color at the (G, GU) coordinate in all panels.

resources are deployed at the target location, which decreases the weight of the Uniform component. More importantly, these opposing effects of metacontrast and attention seem to be operating independently since the weight functions (i.e., the weight of the Uniform component as a function of SOA) undergo vertical shifts with CTOA.

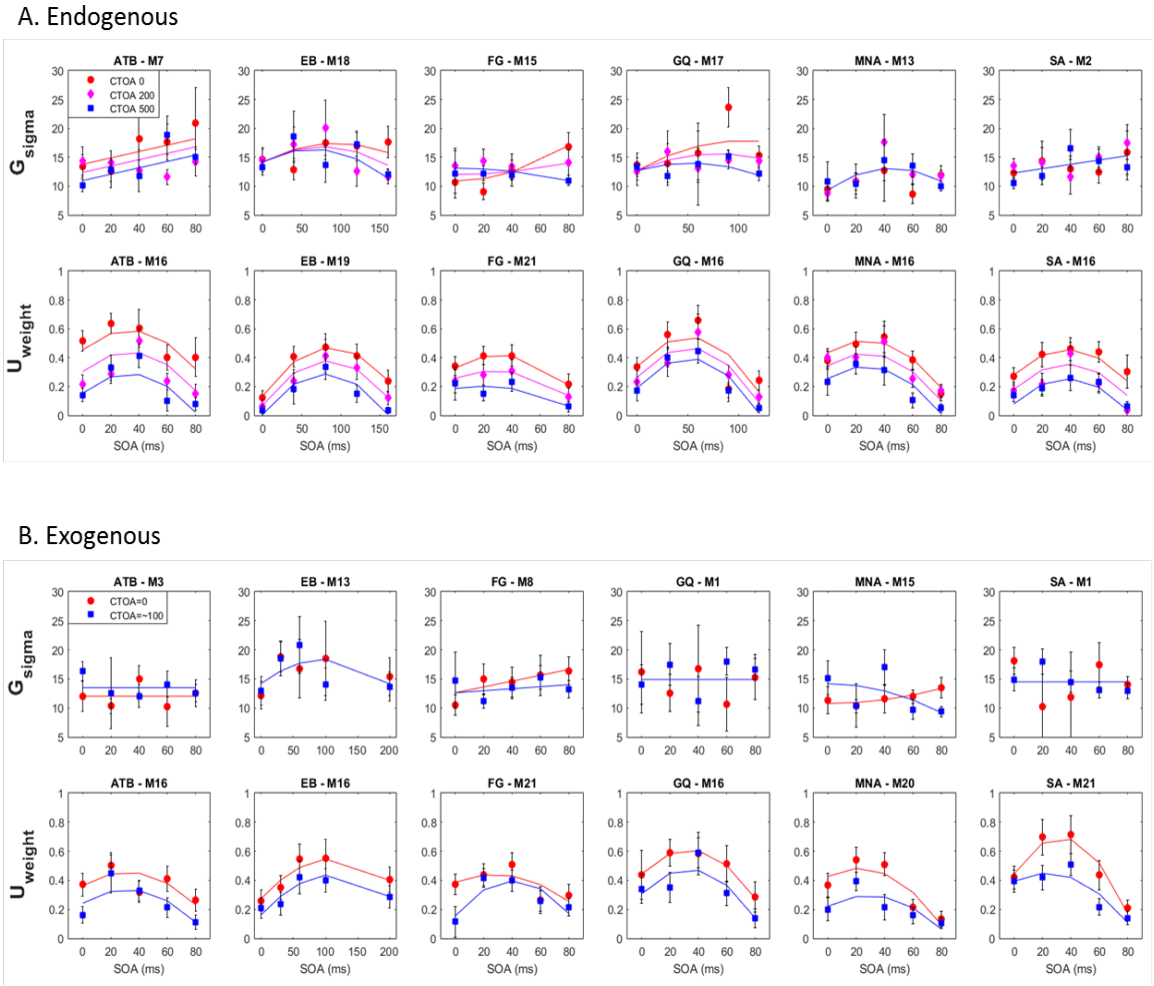

Figure 4-6 The parameters of the winning GU model for all observers in the (A) endogenous and (B) exogenous attention conditions. In each part, first rows represent the standard deviation of the Gaussian whereas the second rows represent the weight of the Uniform component.

These informal evaluations of the results were confirmed by the statistical tests where we fitted model parameters with a series of polynomial regression models. Figure 4-6 also shows the best fitting regression model on top of each panel. For all observers,

the best regression model to capture the changes in the standard deviation of the Gaussian term is different. The inconsistency across observers suggests that masking strength and attentional benefits are not directly reflected in the standard deviation of the Gaussian in the GU model. The changes in the weight of the Uniform term are best captured by the regression model M16 for four out of six observers in the endogenous attention condition. For the observers EB and FG, the best regression model was M19 and M21, respectively. The model M19 has an additional SOAxCTOA interaction term compared to M16, and the model M21 has both SOAxCTOA and SOA^2xCTOA interaction terms (see Table 3-II in Chapter 3 for a complete list of all regression models). In the exogenous attention condition, the best regression model for the weight of the Uniform term was M16 for three out of six observers. Consistently, the best model for observer FG was again M21. However, for observer EB, there was no interaction between SOA and CTOA in the exogenous attention condition. Moreover, for observers MNA and SA, the best regression models were M20 and M21, respectively. Both M20 and M21 contain a quadratic SOA and CTOA interaction, which suggest a masking strength-dependent effect of attention. This is apparent in the SOA-dependent drops in the weight parameter with an increase in CTOA for these observers (Figure 4-6B, U_{weight}). Interestingly, we found such interactions neither for the weight parameter in the endogenous attention condition, nor for the transformed performance in both attention conditions.

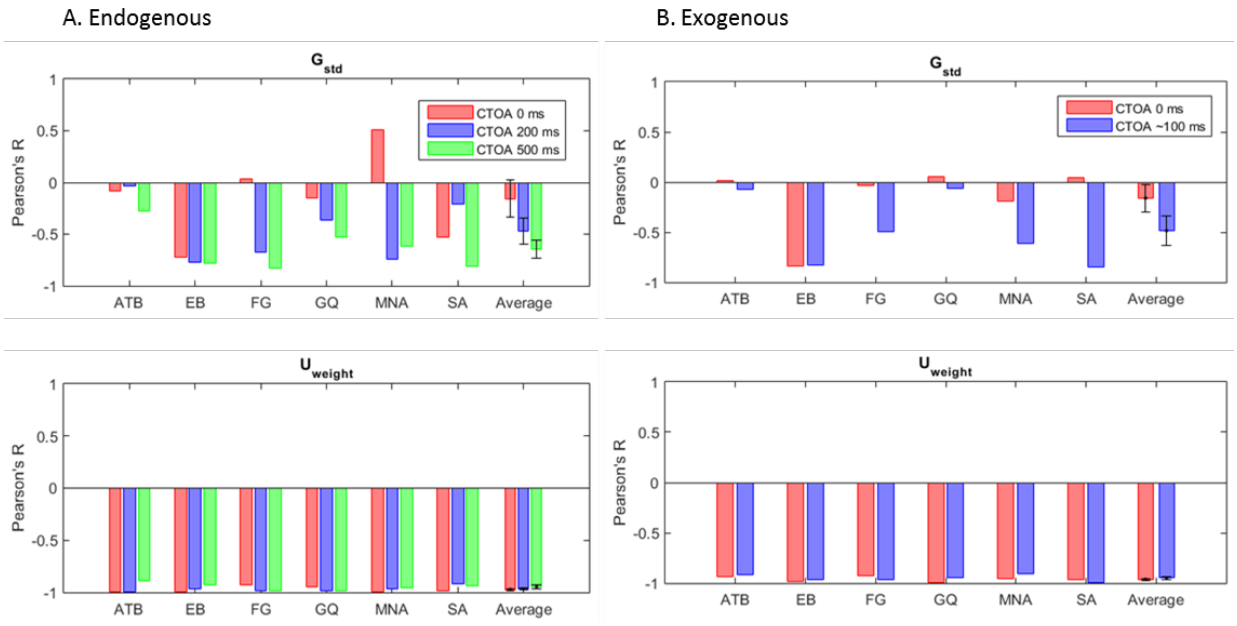

Figure 4-7 The correlations between masking functions and the GU model parameters for the (A) endogenous and (B) exogenous attention conditions. The top row represents the standard deviation of the Gaussian whereas the bottom row represents the weight of the Uniform.

In order to determine how well the changes in transformed performance are reflected in the model parameters, we carried out a correlation analysis, where we computed the Pearson's R coefficient between masking functions and each model parameter separately. Figure 4-7 shows the individual and average correlation coefficients for both attention conditions. Different colors represent different CTOAs. Consistently, we found very strong correlations between the weight of the Uniform and the masking functions in all CTOAs. This suggests that regardless of the level of attentional resources on the target bar, the transformed performance can be closely captured by the changes in the weight of the Uniform term. The standard deviation of the Gaussian term is not correlated with performance when CTOA is zero, and hence, there is no attentional benefits. However, interestingly, in both types of spatial pre-cueing, the correlations tend to increase with increasing CTOA. A one-way repeated measures

ANOVA in the endogenous attention condition did not yield a significant main effect of CTOA on the correlations between masking functions and the standard deviation parameter ($F_{2,10}$=3.824; p=0.058; η_p^2=0.433). However, a paired t-test between CTOA 0 and CTOA 500 ms conditions revealed a significant improvement in correlations (t(5)=2.824, p=0.037). Although there was an increasing trend with CTOA in the exogenous attention condition as well, this difference was not significant (t(5)=2.361, p=0.065).

4.4. Discussion

The visual system is overwhelmed by an enormous amount of information coming through the retina. Since the computing resources available to the brain are limited, they must be used efficiently. Spatial attention comes into play to achieve this feat by selecting a relevant subset of information and filtering out or suppressing the rest. In other words, it controls the quality and quantity of information transfer from sensory input to visual short-term memory (VSTM). Visual masking also plays an important role in the transfer of information from sensory memory to VSTM. In fact, many studies on VSTM have used visual masks to control the information available to the observer, and to isolate processes devoted only to VSTM. However, this approach neglects the possibility of interactions between masking and attention mechanisms. If the mechanisms underlying these two visual phenomena interact, the results of studies where both spatial pre-cues and visual masks are used, need re-interpretation.

Earlier studies on masking and attention relations indeed showed significant interactions between the two (Di Lollo et al., 2000; Enns & Di Lollo, 1997;

Ramachandran & Cobb, 1995; Shelley-Tremblay & Mack, 1999; Smith et al., 2010; Smith & Wolfgang, 2004, 2007; Tata, 2002). In common-onset masking, where the target and mask onsets coincide but the mask outlasts the target, Enns and Di Lollo (1997) showed that attentional benefits due to a spatial pre-cue or reduced set-size strongly depend on mask duration. Similarly, by using metacontrast masks, Tata (2002) showed that increasing set size results in an SOA-dependent impairment of performance. However, most of these studies suffered from ceiling/floor effects and other methodological artefacts (see Discussion in Chapter 3). For instance, in Tata's experiments, there was essentially no masking at all for set size one; percentage of correct responses as a function of SOA formed a flat line at around 95%. However, for larger set sizes, they found strong masking effects, and therefore, this led the author to conclude that attention and metacontrast masking interact.

Evidence from recent studies where these artefacts were avoided, suggests that mechanisms underlying common-onset masking and attention are indeed independent (Argyropoulos et al., 2013; Filmer et al., 2014, 2015; Pilling et al., 2014). Here, we sought to determine whether the same relationship holds for metacontrast masking and attention. In Chapter 3, we varied the set size and SOA to control the attentional load and masking strength, respectively. There are two caveats with this methodology. First, the temporal dynamics of attention and mask interactions cannot be examined by solely manipulating set size. Second, since which one of the oriented bars is the target was unknown to the observers in the beginning of each trial, they had to attend to the entire display. Moreover, the target bar was indicated by the onset of a mask. Therefore, the task employed both endogenous and exogenous attention. Differential contributions from

endogenous and exogenous attention mechanisms might have obscured a potential interaction between metacontrast masking and attention in Chapter 3. Here, we investigated the relationship between metacontrast masking and these two different types of attention separately by using spatial pre-cues. The task of the observers was again to report the orientation of the cued bar. We kept the set size fixed, and varied the CTOA between the pre-cue and the target array and the SOA between the target and mask arrays. We found that for both attention types, the mean absolute errors are affected by CTOA equally at all SOA values. In other words, masking functions underwent vertical shifts with changes in CTOA, indicating that metacontrast masking and attention arise from independent processes. Interestingly, when we further examined the distribution of signed response errors, we found interactions for some of the observers: Two (three) out of six observers in the endogenous (exogenous) attention condition showed significant interactions between CTOA and the guess rate (i.e., the weight of the Uniform distribution). Taken together, these results suggest that although there might be weak interactions between metacontrast masking and attention mechanisms, any joint effect of the two can be well estimated by independent and additive processes.

4.4.1. Implications for Models of Attention

Next, we will discuss whether and how our results can be explained by two prominent models of attention in the literature, namely the Perceptual Template Model (PTM) developed by Lu and Dosher (1998), and the Integrated System Model (ISM) developed by Smith and colleagues (Smith & Wolfgang, 2004 – early version, no explicit VSTM layer; Smith & Ratcliff, 2009 – VSTM stage is added; Smith et al., 2010 – final version). These models are selected since they also address visual masking and its

proposed interactions with attention. In short, PTM can distinguish three attention mechanisms that have distinct signatures on behavioral improvements in perceptual tasks. According to PTM, attention enhances visual stimuli, removes external noise, and reduces multiplicative internal noise. These mechanisms can work in tandem or separately depending on the stimulus configuration and the amount of noise in the stimuli. ISM assumes that attention affects the rate of information transfer from sensory memory to VSTM (Carrasco & McElree, 2001). Masks either truncate sensory information prematurely, before it is fully transferred to VSTM, or add noise to the stimulus, which in turn, slows down the rate at which encoded stimulus information becomes available for later stages of processing (Smith et al., 2010). Moreover, ISM also assumes that masking and attention mechanisms interact, and hence, predicts larger attentional benefits when a stimulus is masked compared to when it is unmasked. Likewise, the stronger the masking is, the larger the attentional effects will be (see Sections 1.5 and 1.6 for a detailed review of these models).

One way masking and attention are related in PTM is that the mask adds noise through temporal integration at the stage of the perceptual template, where stimulus enhancement mechanism of attention also operates. Moreover, in a series of studies, Dosher and Lu showed that external noise exclusion is the mechanism underlying endogenous attention effects whereas both external noise exclusion and stimulus enhancement are in play when exogenous attention operates (Dosher & Lu, 2000a, 2000b; Lu & Dosher, 1998, 2000). Therefore, PTM predicts that the amount of external noise added due to the mask will decrease with increasing SOA, hence a Type-A masking function. PTM further predicts that the effect of attention should be large when external

noise is large compared to the signal. Hence, it predicts that the effect of attention should be largest at SOA=0 and should decrease with increasing SOA. These predictions clearly do not hold for our findings. We obtained Type-B masking functions with an increasing, rather than decreasing masking effects as SOA increases from zero. Furthermore, we found that the effect of both endogenous and exogenous pre-cues, measured by mean absolute errors in orientation judgments, is virtually the same across all SOAs. Based on Type-B shape of masking functions, one could speculate that, by some unspecified mechanism, the metacontrast mask adds external noise in an SOA-dependent manner, i.e., less noise at very short and long SOAs and more noise at intermediate SOAs where masking is strongest. According to this scenario, an increase in CTOA should lead to larger drops in performance at intermediate SOAs compared to short and long SOAs. This is equivalent to a statistical interaction between SOA^2 and CTOA. Although changes in masking functions with CTOA are not in line with this assumption, we found such interactions in statistical models of response errors. Revealed by statistical modeling of the distribution of signed response-errors, rather than just the mean absolute errors, the interaction between SOA^2 and CTOA was evident in the frequency of random guessing behavior for two observers in the endogenous attention condition, and for three observers in the exogenous attention condition. In sum, although the underlying neurophysiological mechanism is unspecified at this time, our finding that there might be modest interactions between metacontrast masking and attention can be explained from PTM.

ISM makes predictions similar to those of PTM. However, as mentioned before, ISM directly incorporates interacting masking and attention mechanisms. For instance, it predicts that there will be no effect of attention in the absence of masks. However, our

baseline data, which correspond to no mask conditions (see Figure 4-3C), show clear effects of attention and we did not find strong evidence in favor of interactions between attention and masking: our masking data showed modest interactions only in the distribution of responses but not in mean performance.

4.4.2. Implications for Masking Models

Attention has facilitative and inhibitory effects in almost all perceptual tasks (Posner, 1980; Smith et al., 2004). However, many early models of visual masking do not address the effects of attention on masking, and mostly assume that attention and masking are independent processes (e.g., Bachmann, 1984; Breitmeyer & Ganz, 1976; Ogmen, 1993). These models can be extended straightforwardly to include attention as an add-on process, which adds to the masking strength or reduces it. Michaels and Turvey's model (1979) included attention as an independent process, which modulates spatial inhibitory projections.

At least one theory of visual masking puts more weight on attention (Di Lollo et al., 2000; Enns & Di Lollo, 1997). In a common onset masking paradigm, Enns and Di Lollo (1997) showed that four-dot masks can produce strong masking when the stimuli were viewed peripherally *and* when attention was diffused to more than one spatial location. Enns and Di Lollo interpreted these effects as a result of higher-level processes of *object substitution*. In summary, interaction between attention and masking is an essential ingredient of the object substitution theory. This prediction was supported by significant interactions found in their study (Di Lollo et al., 2000; Enns & Di Lollo, 1997). However, more recent evidence shows that their results suffered from ceiling/floor artifacts, and that masking and attention do not interact (Argyropoulos et al., 2013;

Filmer et al., 2014; Pilling et al., 2014). Another strong contradiction to the object substitution theory comes from a study by Filmer, Mattingley, and Dux (2015). They found strong masking effects for attended and *foveated* targets. Consistent with these recent reports, here showed that metacontrast masking and attention weakly interact but they can be treated as independent additive processes since this interaction was not apparent in the average performance.

According to the dual-channel models of masking, each stimulus generates a fast transient and a slow sustained activity in separate pathways (channels) (e.g., Breitmeyer & Ogmen, 2006). Visual masking can be explained by inter-channel and/or intra-channel inhibition depending on the stimulus configuration and relative strengths of target and mask. In the context of PTM, inhibition of the target activity due to a mask can be considered as an attenuation after target signal is filtered by the perceptual template, or an increase in the internal additive noise. In other words, rather than adding external noise, metacontrast masking might be operating at the same stage as signal enhancement mechanism of attention. Equivalently, both processes might be modulating the internal additive noise independently. If this assumption holds (i.e., mask does not add external noise but modulate internal additive noise), then the primary attention mechanism at play should be signal enhancement. Our results that attention increases target SNR whereas metacontrast reduces it independently are consistent with this line of reasoning. To be more specific, these findings suggest that metacontrast masking and signal enhancement mechanism of attention do not interact.

It would be interesting to see whether this finding holds for the cases where there is high external noise in the stimuli. For example, noise plus metacontrast masks (i.e.,

compound masks) can be used in conjunction with spatial cues to test whether masking and external noise exclusion mechanism of attention also do not interact.

Chapter 5. Summary and Conclusions

A stimulus (mask) reduces the visibility of another stimulus (target) when they are presented in close spatio-temporal vicinity of each other, a phenomenon called visual masking. Visual masking has been extensively studied to understand dynamics of information processing in the visual system. Another process known to modulate information processing and transfer within the visual system is visual spatial attention. Since both visual masking and visual attention control the transfer of information from sensory memory to VSTM, a natural question is whether these processes interact or operate independently. From a theoretical point of view, determining whether these two processes interact or not can contribute to our understanding of how information is transferred from sensory memory to VSTM. From an empirical point of view, this understanding is especially important when one wants to compare findings from different studies of VSTM, which employ different types of masks or masks with different strengths. If, indeed, masking and attention interact, reconciliation or comparison of findings across different studies will require one to take into account the interaction effects.

Many models have been proposed to explain masking and attention, however, very few of them addresses the relation between the two. Throughout this dissertation, we aimed to first use a statistical modeling technique to assess the distribution of behavioral responses when a target stimulus is masked, and then we sought to determine whether and how masking and attention interact based on this approach.

In Chapter 2, we adopted a statistical point of view, rather than a mechanistic one, to investigate how mask-related activities might influence target-related ones within the context of visual masking. We modeled the distribution of response errors of human observers in three different visual masking paradigms, namely para-/meta-contrast masking, pattern masking by noise, and pattern masking by structure. We adopted statistical models, which have been used previously in studies of visual short-term memory, to capture response characteristics of observers under masking conditions. We tested the following hypotheses (see Specific Aims). *Hypothesis 1:* Mask activity may "interfere" with the encoding of a target and cause decreased precision in observers' reports. *Hypothesis 2:* Mask activity may reduce a target's signal-to-noise ratio (SNR) without interfering with its encoding precision. *Hypothesis 3:* Decreased performance due to masking may result from the confusion or "misbinding" of a mask's features with those of the target, when they are similar as in the case of pattern masking by structure. Among these three hypotheses, our results clearly reject the third hypothesis. Although the evidence from statistical modeling speaks against the second hypothesis, we cannot still reject it due to significant correlations between encoding precision and masking strength. In addition, we found strong support for the second hypothesis. More specifically, our results show that, in all three types of masking, the reduction of a target's SNR (indicated by the increase in guess-rate) was the primary process whereby masking occurred. A significant decrease, which was correlated with masking strength, in the precision of the target's encoding was observed in para-/meta-contrast and pattern masking by structure, but not in pattern masking by noise. We interpret these findings as the mask reducing the target's SNR (i) by suppressing or interrupting the signal of the

target in para-/meta-contrast, (ii) by increasing noise in pattern masking by noise, and (iii) by a combination of the two in pattern masking by structure.

Recent evidence on the relation between masking and attention suggests that studies that reported strong interactions between these two processes suffered from ceiling and/or floor effects. The objective of Chapter 3 and Chapter 4 was to investigate whether metacontrast masking and attention interact by using an experimental design in which saturation effects are avoided. In Chapter 3, we tested the following hypothesis. *Hypothesis 4:* Attention and metacontrast masking are independent processes. We asked observers to report the orientation of a target bar randomly selected from a display containing either 2 or 6 bars. The mask was a ring that surrounded the target bar. Attentional load was controlled by set-size and masking strength by the stimulus onset asynchrony between the target bar and the mask ring. We investigated interactions between masking and attention by analyzing two different aspects of performance: (i) the mean absolute response-errors and (ii) the distribution of signed response-errors. In contrast to previous studies, our results do not reject the null hypothesis (i.e., Hypothesis 4), and show that attention affects observers' performance without interacting with masking. Statistical modeling of response errors suggests that attention and metacontrast masking exert their effects by independently modulating the probability of "guessing" behavior.

Modulating attentional load via varying set size may engage both the voluntary (endogenous) and reflexive (exogenous) mechanisms of attention. In Chapter 4, we used spatial pre-cues to test whether or not there is a significant interaction between masking and endogenous (*Hypothesis 5*) or exogenous (*Hypothesis 6*) attention. Observers were

asked to report the orientation of a bar, which was cued by either a central cue (i.e., an arrow) or a peripheral one (i.e., a small square in the vicinity of the target location). We manipulated the temporal distance between the cue onset and target-array onset (CTOA) to investigate mask-attention relations over time. As expected, the effect of attention was apparent at different CTOAs for endogenous and exogenous attention. However, our results show that endogenous and exogenous attention cannot be dissociated solely based on mean absolute errors or distribution of signed response-errors. Moreover, our results accept the null hypotheses expressed above, and once again suggest that metacontrast masking and attention do not interact at the level of mean absolute response errors.

Overall, these results add to the growing body of evidence that metacontrast masking and attention do not interact, at least in a way that it would require re-interpretation of studies that employed various types of masks and spatial cues. In other words, metacontrast masking and attention can be safely considered as mostly independent processes, and their joint effects can be well estimated by a simple addition of their individual effects on a perceptual task. Nevertheless, both visual masking and attention are not unitary processes. For instance, masking of surface properties is different from masking of contour properties. Further, the same level of reduction in visibility can be obtained through completely different neural mechanisms under different stimuli configurations and tasks. Similarly, computational, psychophysical, and neuroimaging studies show that attentional effects might stem from distinct mechanisms depending on the task demands and spatiotemporal properties of the stimuli. Therefore, it is necessary to determine the relationship between masking and attention in different contexts and criterion contents, and with various tasks.

Chapter 6. Future Directions

Visual masking and attention are not unitary phenomena. Their effects depend on various spatiotemporal parameters of the stimuli as well as on cognitive requirements of the task. Early studies showed significant interactions between two types of masking (common-onset masking and metacontrast masking) and attention. However, most of these studies suffered from ceiling/floor artefacts. Recent studies on common-onset masking revealed the absence of interaction between masking and attention. In this dissertation, we showed that the same is true for metacontrast masking and attention. However, it remains to be established whether this relationship holds for other masking types and attention. Feature-based and object-based attentional benefits are also well established. The underlying mechanisms of these processes differ at least partially from those of spatial attention. Therefore, how masking and feature-based or object-based attention interact also need to be investigated.

Earlier studies, which showed significant interactions between attention and masking employed either simple detection tasks or easy discrimination tasks. However, in all experiments in this dissertation, the task was to report the orientation of a target bar as accurately as possible (with steps of 1 degree of angle). This can be categorized as a difficult discrimination task. The effects of attention and masking are known to depend on task difficulty. Simple detection tasks for instance are limited by luminance contrast whereas fine-grained discrimination tasks require information about multiple attributes of the stimuli since it depends on stimulus similarity. It is well known that the magnocellular pathway and its associated transient mechanisms have very different

contrast responses compared to parvocellular pathway and its associated sustained mechanisms (Croner & Kaplan, 1995; Kaplan & Shapley, 1986). Simple detection and easy discrimination tasks can be carried out by both transient and sustained mechanisms, whereas difficult fine-discrimination tasks are likely to necessitate sustained mechanisms. Therefore, it would be informative to see if interaction between masking and attention depends on task requirements. Recognition based on conjunction of multiple features in the presence of masks needs to be investigated under different levels of task difficulty, (i.e., by modulating the number of features required to uniquely describe the target item) with and without deploying attentional resources on the locus of the target location.

The current models of attention, which also address its relation to masking, inherently assume that there are strong interactions between these two processes. In light of the evidence presented here and recent studies on common-onset masking, these models need revisions. It would be interesting to investigate whether and how these models (e.g., PTM by Dosher, Lu, and colleagues, and ISM by Smith and colleagues) can be updated to accommodate these recent findings.

Models of visual masking have mostly remained agnostic as to how attention and masking work in tandem. However, this does not necessarily mean that these models argue against any potential modulation of masking effect by attention. Attention can be added as an add-on process to most of these models. In fact, Michaels and Turvey's model employs attention as an independent add-on process. One of the most comprehensive and detailed model of masking is the dual-channel model (Breitmeyer & Ganz, 1976; Ogmen, 1993). This model is based on empirical findings on characteristics of different types of cells in the retina and throughout the visual pathways. Since our

results suggest that masking and attention mostly do not interact, it would be interesting to see how attention can be added to this neural network model of masking. Depending on which level(s) of processing in the visual processing hierarchy attention can modulate, the dual-channel model predicts different outcomes. For instance, if attention modulates only the sustained channel (or the transient channel), its effect will be different for backward and forward masking. From this point of view, attention might be interacting with forward masking but not backward masking. These points need further experimentation and quantitative modeling.

Finally, neuroimaging studies are needed to determine whether the lack of interaction between masking and attention is also evident in the functionally active networks in the brain. It might be that attention and masking might arise from completely independent or overlapping networks. In the latter case, careful analysis of single cell recordings or event-related potentials to delineate the temporal dynamics of mask-attention relations within a particular cortical region, would be much needed. For instance, attention is shown to sharpen the orientation tuning and increase contrast gain of some cortical cells. By measuring firing rates as a function of time while manipulating attention by spatial cues and modulating mask strength, one can distinguish the effects of masking and attention, and whether their effects interact at the neural level, or not.

References

Ağaoğlu, M. N., Herzog, M. H., & Ogmen, H. (2012). Non-retinotopic feature processing in the absence of retinotopic spatial layout and the construction of perceptual space from motion. *Vision Research*, *71*, 10–17.

Agaoglu, S., Agaoglu, M. N., Breitmeyer, B. G., & Ogmen, H. (2015). A statistical perspective to visual masking. *Vision Research*, *115*, 23–39.

Ansorge, U., Klotz, W., & Neumann, O. (1998). Manual and verbal responses to completely masked (unreportable) stimuli: exploring some conditions for the metacontrast dissociation. *Perception*, *27*(10), 1177–1189.

Argyropoulos, I., Gellatly, A., Pilling, M., & Carter, W. (2013). Set size and mask duration do not interact in object-substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 646–661.

Atkinson, R. C., & Shiffrin, R. M. (1971). The Control of Short-Term Memory. *Scientific American*, *225*(2), 82–90.

Averbach, E., & Coriell, A. S. (1961). Short-term momory in vision. *Bell System Technical Journal*, *40*(1), 309–328.

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

Bachman, T. (1994). *Psychophysiology of visual masking: the fine structure of conscious experience*. Commack, NY: Nova Science.

Bachmann, T. (1984). The process of perceptual retouch: Nonspecific afferent activation dynamics in explaining visual masking. *Perception & Psychophysics*, *35*(1), 69–84.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 1–11.

Boyer, J., & Ro, T. (2007). Attention attenuates metacontrast masking. *Cognition*, *104*, 135–149.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, *83*(1), 1–36.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*(8), 1572–1595.

Breitmeyer, B. G., & Ogmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision* (2nd Ed.). Oxford: Oxford University Press.

Briand, K. A., & Klein, R. M. (1987). Is Posner's "beam" the same as Treisman's "glue"?: On the relation between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 228–241.

Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychological Review*, *78*(6), 528–539.

Bridgeman, B. (1978). Distributed sensory coding applied to simulations of iconic storage and metacontrast. *Bulletin of Mathematical Biology*, *40*(5), 605–623.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722.

Burr, D. (1980). Motion Smear. *Nature*, *284*, 164–165.

Burr, D. (2004). Eye movements: keeping vision stable. *Current Biology*, *14*, R195–R197.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review*, *101*(3), 446–469.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.

Carrasco, M., & Barbot, A. (2015). How Attention Affects Spatial Resolution. *Cold Spring Harbor Symposia on Quantitative Biology*.

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(9), 5363–5367.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research*, *40*, 1203–1215.

Cheal, M., & Lyon, D. R. (1991). Importance of precue location in directing attention. *Acta Psychologica*, *76*(3), 201–211.

Chen, S., Bedell, H. E., & Ogmen, H. (1995). A target in real motion appears blurred in the absence of other proximal moving targets. *Vision Research*, *35*(16), 2315–2328.

Chen, Y., Martinez-Conde, S., Macknik, S. L., Bereshpolova, Y., Swadlow, H. A., & Alonso, J.-M. (2008). Task difficulty modulates activity of specific neuronal populations in primary visual cortex. *Nature Neuroscience*, *11*(8), 974–982.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, *27*(3), 183–228.

Cowan, N. (2000). The magical number 4 in short-term memory : A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114 discussion 114–185.

Cowan, N. (2005). *Working Memory Capacity*. New York: Psychology Press.

Cowan, N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, *19*(1), 51–57.

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research*, *35*(1), 7–24.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*(4), 481–507.

Dombrowe, I., Hermens, F., Francis, G., & Herzog, M. H. (2009). The roles of mask luminance and perceptual grouping in visual backward masking. *Journal of Vision*, *9*(11), 1–11.

Dosher, B. A., & Lu, Z.-L. (2000a). Mechanisms of perceptual attention in precuing of location. *Vision Research*, *40*, 1269–1292.

Dosher, B. A., & Lu, Z.-L. (2000b). Noise Exclusion in Spatial Attention. *Psychological Science*, *11*(2), 139–146.

Ebbeler, D. H. (1975). On the probability of correct model selection using the maximum R2 choise criterion. *International Economic Review*, *16*(2), 516–520.

Eckstein, M. P. (1998). The Lower Visual Search Efficiency for Conjunctions Is Due to Noise and not Serial Attentional Processing. *Psychological Science*, *9*(2), 111–118.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics*, *62*(3), 425–451.

Egeth, H. E., & Yantis, S. (1997). Visual attention: control, representation, and time course. *Annual Review of Psychology*, *48*, 269–297.

Enns, J. T. (2002). Visual binding in the standing wave illusion. *Psychonomic Bulletin & Review*, *9*(3), 489–496.

Enns, J. T., & Di Lollo, V. (1997). Object Substitution: A new form of masking in unattended visual locations. *Psychological Science*, *8*(2), 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352.

Ester, E. F., Zilber, E., & Serences, J. T. (2015). Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of Vision*, *15*(1), 1–12.

Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2014). Size (mostly) doesn't matter: the role of set size in object substitution masking. *Attention, Perception & Psychophysics*, *76*, 1620–1629.

Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2015). Object substitution masking for an attended and foveated target. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 6–10.

Francis, G. (2000). Quantitative Theories of Metacontrast Masking. *Psychological Review*, *107*(4), 768–785.

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–182.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135.

Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 845–866.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, *1*(1), 17–61.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences*, *6*, 1–54.

Haber, R. N., & Standing, L. G. (1970). Direct estimates of the apparent duration of a flash. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *24*(4), 216–229.

Haenny, P. E., Maunsell, J. H. R., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Experimental Brain Research*, *69*(2), 245–259.

Harrison, G. W., Rajsic, J., & Wilson, D. E. (2014). Effect of object-substitution masking on the perceptual quality of object representations. *Journal of Vision*, *14*(10).

Herrmann, K., Montaser-Kouhsari, L., Carrasco, M., & Heeger, D. J. (2010). When size matters: attention affects performance by contrast or response gain. *Nature Neuroscience*, *13*(12), 1554–1559.

Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: feature inheritance and shine-through. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 4271–4275.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research*, *33*(9), 1219–1240.

Hirose, N., & Osaka, N. (2010). Asymmetry in object substitution masking occurs relative to the direction of spatial attention shift. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(1), 25–37.

Hocking, R. R. (1976). A Biometrics invited paper. The analysis and selection of variables in linear regression. *Biometrics*, *32*(1), 1–49.

Hofer, D., Walder, F., & Groner, M. (1989). Metakontrast: ein berühmtes, aber schwer messbares Phänomen. *Schweizerische Zeitschrift für Psychologie und ihre Anwendungen*, *48*, 219–232.

Jeffreys, H. (1998). *The theory of probability*. Oxford University Press.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (Vol. 9, pp. 187–203). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 2755–2757.

Kay, S. (2005). Exponentially embedded families-new approaches to model order estimation. *IEEE Transactions on Aerospace and Electronic Systems*, *41*(1), 333–345.

Klotz, W., & Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 976–992.

Kröse, B. J. A., & Julesz, B. (1989). The control and speed of shifts of attention. *Vision Research*, *29*(11), 1607–1619.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*(4), 375–381.

Ling, S., & Carrasco, M. (2006). When sustained attention impairs perception. *Nature Neuroscience*, *9*(10), 1243–1245.

Loftus, G. R., Duncan, J., & Gehrig, P. (1992). On the Time Course of Perceptual Information That Results From a Brief Visual Presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 530–549.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes mechanisms of attention. *Vision Research*, *38*(9), 1183–1198.

Lu, Z.-L., & Dosher, B. A. (2000). Spatial Attention : Different Mechanisms for Central and Peripheral Temporal Precues? *Journal of Experimental Psychology*, *26*(5), 1534–1548.

Lu, Z.-L., & Dosher, B. A. (2004). Spatial attention excludes external noise without changing the spatial frequency tuning of the perceptual template. *Journal of Vision*, *4*, 955–966.

Lu, Z.-L., & Dosher, B. A. (2005). External noise distinguishes mechanisms of attention. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 448–453). Burlington: Academic Press.

Lu, Z.-L., Jeon, S. T., & Dosher, B. A. (2004). Temporal tuning characteristics of the perceptual template and endogenous cuing of spatial attention. *Vision Research*, *44*, 1333–1350.

Lu, Z.-L., Lesmes, L. A., & Dosher, B. A. (2002). Spatial attention excludes external noise at the target location. *Journal of Vision*, *2*, 312–323.

Lu, Z.-L., Liu, C. Q., & Dosher, B. A. (2000). Attention mechanisms for multi-location first- and second-order motion perception. *Vision Research*, *40*(2), 173–186.

Mackay, D. J. C. (2004). *Information Theory, Inference, and Learning Algorithms* (7th ed.). Cambridge University Press.

Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review*, *14*(6), 1072–1078.

Michaels, C. F., & Turvey, M. T. (1979). Central sources of visual masking: Indexing structures supporting seeing at a single, brief glance. *Psychological Research*, *41*, 1–61.

Montagna, B., Pestilli, F., & Carrasco, M. (2009). Attention trades off spatial acuity. *Vision Research*, *49*(7), 735–745.

Montaser-Kouhsari, L., & Carrasco, M. (2009). Perceptual asymmetries are preserved in short-term memory tasks. *Attention, Perception & Psychophysics*, *71*(8), 1782–1792.

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology. Human Perception and Performance*, *15*(2), 315–30.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*(11), 1631–1647.

Noory, B., Herzog, M. H., & Ogmen, H. (2015). Spatial properties of non-retinotopic reference frames in human vision. *Journal of Vision*, *113*, 44–54.

Ogmen, H. (1993). A Neural Theory of Retino-Cortical Dynamics. *Neural Networks*, *6*, 245–273.

Ogmen, H., Breitmeyer, B. G., & Melvin, R. (2003). The what and where in visual masking. *Vision Research*, *43*(12), 1337–1350.

Ogmen, H., Ekiz, O., Huynh, D., Bedell, H. E., & Tripathy, S. P. (2013). Bottlenecks of motion processing during a visual glance: the leaky flask model. *PLOS One*, *8*(12), 1–12.

Ogmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, *46*, 3234–3242.

Otto, T. U., Ogmen, H., & Herzog, M. H. (2006). The flight path of the phoenix - The visible trace of invisible elements in human vision. *Journal of Vision*, *6*, 1079–1086.

Otto, T. U., Ogmen, H., & Herzog, M. H. (2008). Assessing the microstructure of motion correspondences with non-retinotopic feature attribution. *Journal of Vision*, *8*(7), 1–15.

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 332–350.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*(13), 1703–1721.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(1), 108–130.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*, 1227–1268.

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*, 97–107.

Peelen, M. V, Heslenfeld, D. J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *NeuroImage*, *22*, 822–830.

Pestilli, F., & Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision Research*, *45*(14), 1867–1875.

Pestilli, F., Ling, S., & Carrasco, M. (2009). A population-coding model of attention's influence on contrast response: Estimating neural effects from psychophysical data. *Vision Research*, *49*, 1144–1153.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290.

Pilling, M., Gellatly, A., Argyropoulos, Y., & Skarratt, P. (2014). Exogenous spatial precuing reliably modulates object processing but not object substitution masking. *Attention, Perception & Psychophysics*, *76*(6), 1560–1576.

Polk, T. A., Drake, R. M., Jonides, J. J., Smith, M. R., & Smith, E. E. (2008). Attention enhances the neural processing of relevant features and suppresses the processing of irrelevant features in humans: a functional magnetic resonance imaging study of the Stroop task. *The Journal of Neuroscience*, *28*(51), 13786–13792.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: the role of set for spatial location. In H. L. J. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hilldale, NJ: Lawrence Erlbaum Associates.

Purushothaman, G., Ogmen, H., Chen, S., & Bedell, H. E. (1998). Motion deblurring in a neural network model of retino-cortical dynamics. *Vision Research*, *38*, 1827–1842.

Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature*, *373*, 66–68.

Raymond, D. (1910). The "pendular whiplash illusion." *Psychonomic Bulletin*, *7*(11).

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics*, *51*(3), 279–290.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience*, *27*, 611–647.

Reynolds, J. H., & Heeger, D. J. (2009). The Normalization Model of Attention. *Neuron*, *61*, 168–185.

Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, *26*, 703–714.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

Sayim, B., Manassi, M., & Herzog, M. H. (2014). How color, regularity, and good Gestalt determine backward masking. *Journal of Vision*, *14*(7), 1–11.

Schill, K., & Zetzsche, C. (1995). A model of visual spatio-temporal memory: the icon revisited. *Psychological Research*, *57*, 88–102.

Schmidt, T. (2002). The finger in flight: real-time motor control by visually masked color stimuli. *Psychological Science*, *13*(2), 112–118.

Shelley-Tremblay, J., & Mack, A. (1999). Metacontrast Masking and Attention. *Psychological Science*, *10*(6), 508–515.

Shiu, L., & Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1037–1054.

Shooner, C., Tripathy, S. P., Bedell, H. E., & Ogmen, H. (2010). High-capacity, transient retention of direction-of-motion information for multiple moving objects. *Journal of Vision*, *10*(6), 1–20.

Smith, P. L. (2000). Attention and luminance detection: effects of cues, masks, and pedestals. *Journal of Experimental Psychology. Human Perception and Performance*, *26*(4), 1401–1420.

Smith, P. L., Ellis, R., Sewell, D. K., & Wolfgang, B. J. (2010). Cued detection with compound integration-interruption masks reveals multiple attentional mechanisms. *Journal of Vision*, *10*(5), 1–10.

Smith, P. L., Lee, Y.-E., Wolfgang, B. J., & Ratcliff, R. (2009). Attention and the detection of masked radial frequency patterns: Data and model. *Vision Research*, *49*, 1363–1377.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, *116*(2), 283–317.

Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision Research*, *44*, 1297–1320.

Smith, P. L., & Wolfgang, B. J. (2004). The attentional dynamics of masked detection. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(1), 119–136.

Smith, P. L., & Wolfgang, B. J. (2007). Attentional mechanisms in visual signal detection: the effects of simultaneous and delayed noise and pattern masks. *Perception & Psychophysics*, *69*(7), 1093–1104.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General & Applied*, *74*(11, Whole No. 498), 1–29.

Sperling, G., & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America*, *58*(8), 1133–1145.

Sreenivasan, K. K., & Jha, A. P. (2007). Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *Journal of Cognitive Neuroscience*, *19*(1), 32–41.

Stewart, A. L., & Purcell, D. G. (1970). U-shaped masking functions in visual backward masking: Effects of target configuration and retinal position. *Perception & Psychophysics*, *7*(4), 253–256.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: an attentional repulsion effect. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(2), 443–463.

Tata, M. S. (2002). Attend to it now or lose it forever: Selective attention, metacontrast masking, and object substitution. *Perception & Psychophysics*, *64*(7), 1028–1038.

Tombu, M. N., Asplund, C. L., Dux, P. E., Godwin, D., Martin, J. W., & Marois, R. (2011). A Unified attentional bottleneck in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13426–13431.

Van den Berg, R., Johnson, A., Anton, A. M., Schepers, A. L., & Cornelissen, F. W. (2012). Comparing crowding in human and ideal observers. *Journal of Vision*, *12*(8), 1–15.

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(22), 8780–8785.

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(10), 6275–6280.

Ward, L. M. (2008). Attention. *Scholarpedia*, *3*(10), 1538.

Wasserman, L. (2000). Bayesian Model Selection and Model Averaging. *Journal of Mathematical Psychology*, *44*, 92–107.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, *238*(4828), 778–780.

Weisstein, N. (1968). A Rashevsky-Landahl neural net - Simulation of Metacontrast. *PSYCHOLOGICAL REVIEW*, *75*(6), 494–521.

Weisstein, N. (1972). Metacontrast. In D. Jameson & L. M. Hurvich (Eds.), *Handbook of Sensory Physiology* (Vol. 7 / 4, pp. 233–272). Berlin, Heidelberg: Springer Berlin Heidelberg.

Weisstein, N., Ozog, G., & Szoc, R. (1975). A comparison and elaboration of two models of metacontrast. *Psychological Review*, *82*, 325–343.

Werner, H. (1935). Studies on contour: I. qualitative analyses. *The American Journal of Psychology*, *47*, 40–64.

Williams, A., & Weisstein, N. (1978). Line segments are perceived better in a coherent context than alone: an object-line effect in visual perception. *Memory & Cognition*, *6*(2), 85–90.

Williams, M. C., & Weisstein, N. (1981). Spatial frequency response and perceived depth in the time-course of object superiority. *Vision Research*, *21*(5), 631–646.

Wilson, A. E., & Johnson, R. M. (1985). Transposition in backward masking. The case of the travelling gap. *Vision Research*, *25*(2), 283–288.

Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, *13*, 55–80.

Wright, R. D., & Ward, L. M. (2008). *Orienting of attention.* New York: Oxford University Press.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, *396*, 72–75.

Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science*, *14*(3), 225–231.

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision*, *10*(10), 1–12.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233–235.

Appendices

Appendix A

Here we demonstrate the robustness of our results by employing another model fitting approach and model comparison metric. We fitted each model by using the LMS technique, which minimizes the squared sum of errors between the model prediction and the actual data. We, then, evaluated the performance of each model by using adjusted R^2 coefficients. Adjusted R^2 is a standard unbiased model selection criterion and takes the number of parameters and the number of samples into account so that models with varying number of parameters can be pitted against each other (Ebbeler, 1975 review: Hocking, 1976). Calculation of adjusted R^2 in terms of coefficient of determination (R^2) is given in Equation (A-1).

$$Adjusted\ R^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}\ , \qquad \text{(A-1)}$$

where n is the sample size (100 trials per SOA) and p is the number of parameters in the model.

Due to circular nature of the error values (-90 to 90 deg), using a wrapped Gaussian (von Mises distribution) function is more appropriate than using a standard Gaussian function. However, when the standard deviation of the Gaussian is small, the difference between single and wrapped Gaussians is negligible (Shooner, Tripathy, Bedell, & Ogmen, 2010). The Probability density function of a wrapped Gaussian in the context of this study can be expressed as

$$f(\varepsilon, \mu, \sigma) = \frac{1}{\gamma} \sum_{k=-\infty}^{\infty} \exp \frac{-(\varepsilon - \mu + k180)^2}{2\sigma^2}, \quad \text{(A-2)}$$

where ε is response error, μ and σ are mean and standard deviation of a standard Gaussian, k represents the number of wrappings, and γ is a normalization factor. When k=0, Equation (A-2) becomes equivalent to a standard Gaussian. The operation described by Equation (A-2) is nothing but adding shifted Gaussians centered on multiples of $\pm$180 deg and normalizing such that area under the curve within $\pm$90 deg error values summed to unity. We picked two different wrappings (k=0: standard Gaussian and k=10) to test if wrappings have any effect on adjusted R^2.

Given that the optimization algorithm can be trapped in a local minimum due to multi-dimensionality of the parameter space (2-, 3-, and 6-dimensional for the G, GU, and Misbinding models, respectively), we ran the model simulations 200 times for each number of wrappings with randomly chosen initial parameters. We picked the optimum parameters, which resulted in the largest adjusted R^2 coefficients, because the larger the adjusted R^2 coefficients are, the better the model performs.

After determining the best model based on adjusted R^2 metric, we quantified the correlation between the model parameters (mean and standard deviation of the Gaussian term, the weight of the Uniform component) and the masking strengths by calculating Pearson R coefficients. A strong correlation between a parameter and the masking strength would suggest a critical role for this parameter for explaining how masking occurs.

Para- /Meta-contrast Masking

The GU model performs better than the G model according to adjusted R^2 values (Figure A-1), consistent with model selection by using BMC. The mean of the Gaussian, which represents bias in orientation judgments, if any, shows no systematic pattern of change as a function of SOA as revealed by a one-way ANOVA ($F(11,44)=0.914$, $p=0.536$, $\eta_p^2=0.186$) (Figure A-2A). However, the standard deviation of the Gaussian in the GU model increases as SOA values approach 50ms (where masking is most effective) and then decreases to a plateau (Figure A-2B). A one-way ANOVA confirms a significant effect of SOA on standard deviation ($F(11,44)=4.056$, $p<0.001$, $\eta_p^2=0.504$). Last but not the least, the weight of the uniform distribution shows significant change with SOA: $F(11,44)=13.601$, $p<0.001$, $\eta_p^2=0.773$) (Figure A-2C). Visual comparison of model coefficients (Figure A-2A - A-2C) with masking strengths (Figure A-2D) reveals that the mean of the Gaussian term does not correlate with masking strength but the standard deviation of the Gaussian term and the weight of the Uniform term do, the latter having stronger correlation than the former. Pearson's R coefficients confirm these qualitative observations. Figure A-2E depicts Pearson's R coefficients averaged across observers for each model parameter. R coefficients for mean did not reach significance ($t(4)=0.310$, $p=0.772$) whereas those for standard deviation and weight were highly significant (standard deviation: $t(4)=10.150$, $p<0.001$; weight: $t(4)=73.722$, $p<0.00001$). Furthermore, the weight of the Uniform was more strongly correlated with the masking strength than the standard deviation of the Gaussian (paired t-test: $t(4)=6.104$, $p=0.004$).

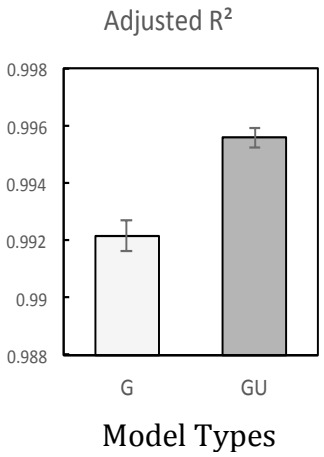

Figure A-1 Adjusted R^2 values obtained from G and GU models averaged across all observers. G represents the Gaussian model whereas GU stands for the Gaussian+Uniform model. Error bars in the average data represent ±SEM, n=5.

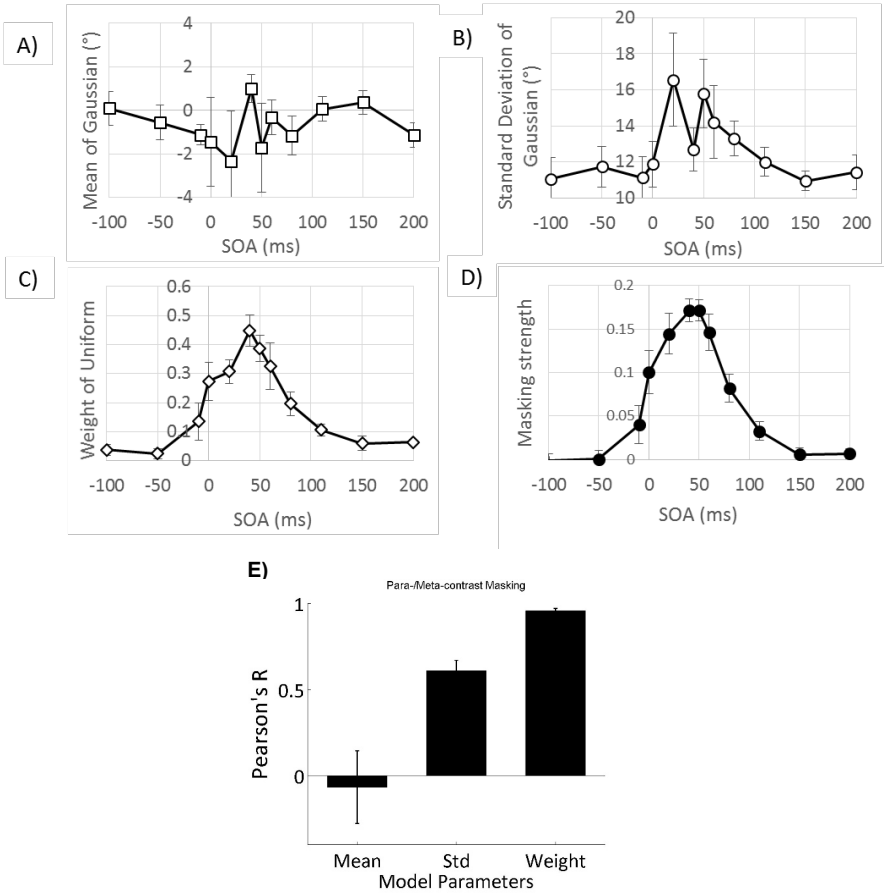

Figure A-2 Parameters of the winning model (GU) for para-/meta-contrast masking. A) Mean of the Gaussian. B) Standard deviation of the Gaussian. C) Weight of the Uniform. D) Masking strengths. E) Correlations between model parameters and masking strengths.

Pattern Masking by Noise

The GU model is the winning model also for pattern masking by noise as indicated by the adjusted R2 values (Figure A-3). Figure A-4 shows model parameters against SOA values. Means of the Gaussian in the GU model again neither show a systematic change with SOA ($F(11,44)=0.615$, $p=0.806$, $\eta_p^2=0.133$) nor significant correlation with masking strength ($t(4)=1.111$, $p=0.329$) (Figure A-4A, A-4D-E). Standard deviation of the Gaussian in the GU model shows changes that resemble the changes in the masking strength (Figure A-4B and A-4D) and there is indeed a relatively weak but significant quadratic trend ($F(1,4)=8.805$, $p=0.041$, $\eta_p^2=0.688$). However, only two of the five observers show this trend and a one-way ANOVA of standard deviation yielded no significant effect of SOA ($F(11,44)=1.350$, $p=0.231$, $\eta_p^2=0.252$).. More importantly, Pearson's R did not differ significantly from zero ($t(4)=1.968$, $p=0.121$) indicating a rather poor correlation between changes in standard deviation and masking strength. In contrast, as with para-/meta-contrast masking, guess rate correlates strongly with the masking strength (Figure A-4C – A-4E): The stronger the effect, the higher the guess rate, reflected in the weight of the uniform component in the GU model (Figure A-4C – A-4D). This SOA-dependent modulation is significant ($F(11,44)=17.764$, $p<0.001$, $\eta_p^2=0.773$). Correlation of the weights with the masking strength was also highly significant ($t(4)=33.140$, $p<0.00001$).

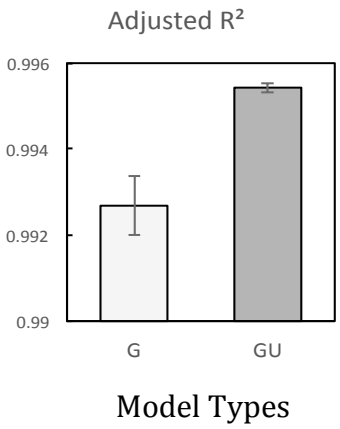

Figure A-3 Adjusted R^2 values obtained from G and GU models averaged across all observers for pattern masking by noise. G represents the Gaussian model whereas GU stands for the Gaussian+Uniform model. Error bars in the average data represent ±SEM, n=5.

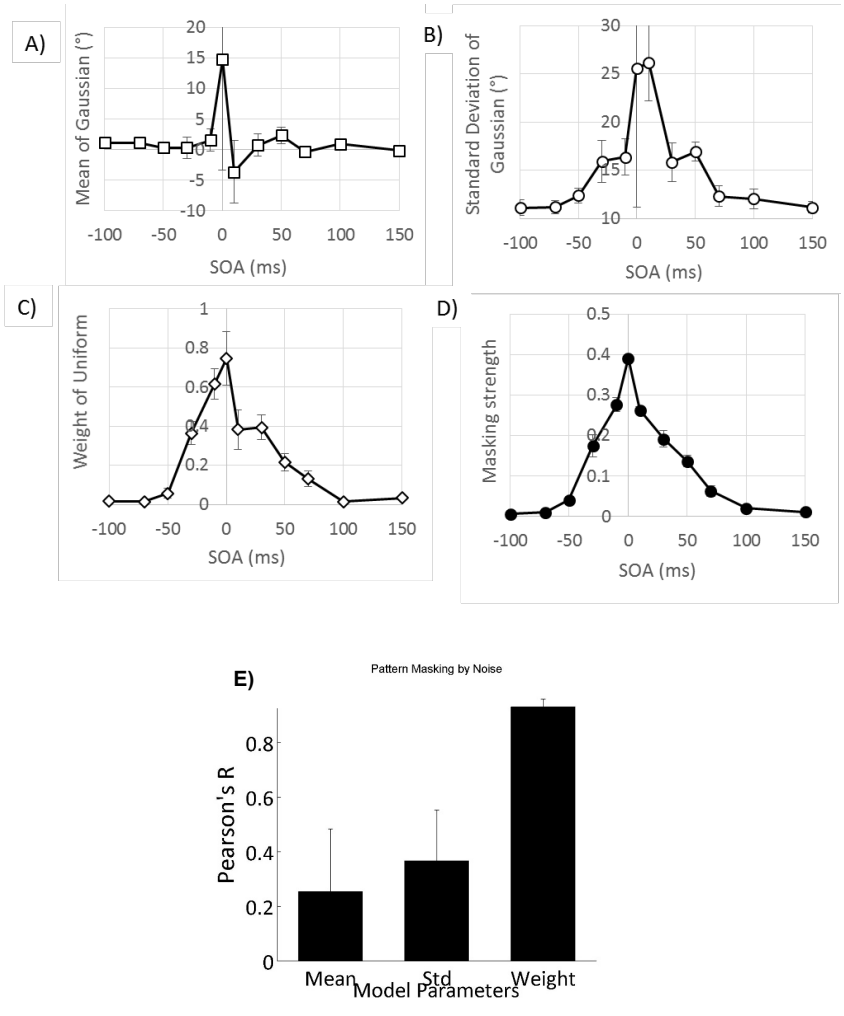

Figure A-4 Parameters of the winning model (GU) in pattern masking by noise. A) Means, and B) Standard deviations of the Gaussian. C) Weight of the Uniform. D) Masking strengths. E) Correlations between model parameters and masking strengths. Error bars represent ±SEM, n=5.

Pattern Masking by Structure

Model performance is shown in Figure A-5. Once again, the GU model outperforms all other models. The mean and standard deviation of the Gaussian term and weight of the uniform distribution in the GU model are plotted against SOAs in Figure A-6A – A-6C. Consistently, we observe neither any systematic change in the means with SOA ($F(1,4)=0.930$, $p=0.521$, $\eta_p^2=0.189$) nor strong correlation with the masking strength (Pearson's R = -0.237±0.438), ruling out the possibility of any orientation bias and any relation to a masking effect. We found a significant effect of SOA on the standard deviation ($F(1,4)=2.640$, $p=0.011$, $\eta_p^2=0.398$). Finally, the weight of the uniform distribution also changed significantly with SOAs ($F(1,4)=13.140$, $p<0.001$, $\eta_p^2=0.767$). The changes in both the standard deviation and the weight of the uniform distribution with SOA correlate well with those of the masking strength (Figure A-6B – A-6E) (Pearson's R for the standard deviation of the Gaussian = 0.584±0.062; for the weight of the Uniform = 0.923±0.062). However, it should be noted that a major factor in producing a masking effect is probably due to a reduction in SNR because the weights of the Uniform term more strictly follow the masking strengths (paired t-test between Pearson's R for the standard deviation of the Gaussian term and the weight of the Uniform: $t(4)=7.874$, $p=0.001$). Therefore, these findings suggest that pattern masking by structure also occurs, from a statistical point of view, primarily due to the reduction of target SNR and only partly due to the interference of the target signal with the mask related activity.

In summary, regardless of whether we take the LMS+adjusted R^2 or the MLE+BMC approach, the results are almost identical, confirming that our findings are robust and not prone to error due to limitations of the methodology used.

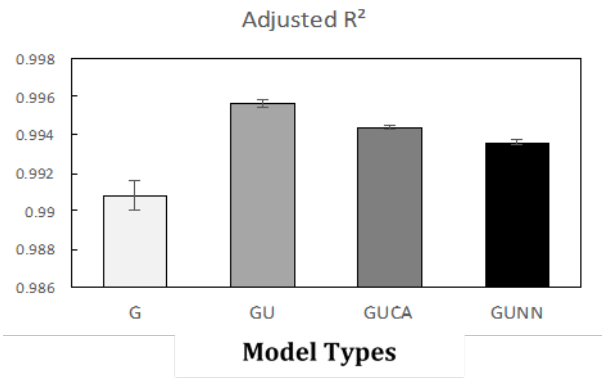

Figure A-5 Adjusted R^2 values obtained from G (Gaussian), GU (Gaussian+Uniform), GUCA (Misbinding – Closest Angle), and GUNN (Misbinding – Nearest Neighbor) models in pattern masking by structure. Error bars represent ±SEM, n=5.

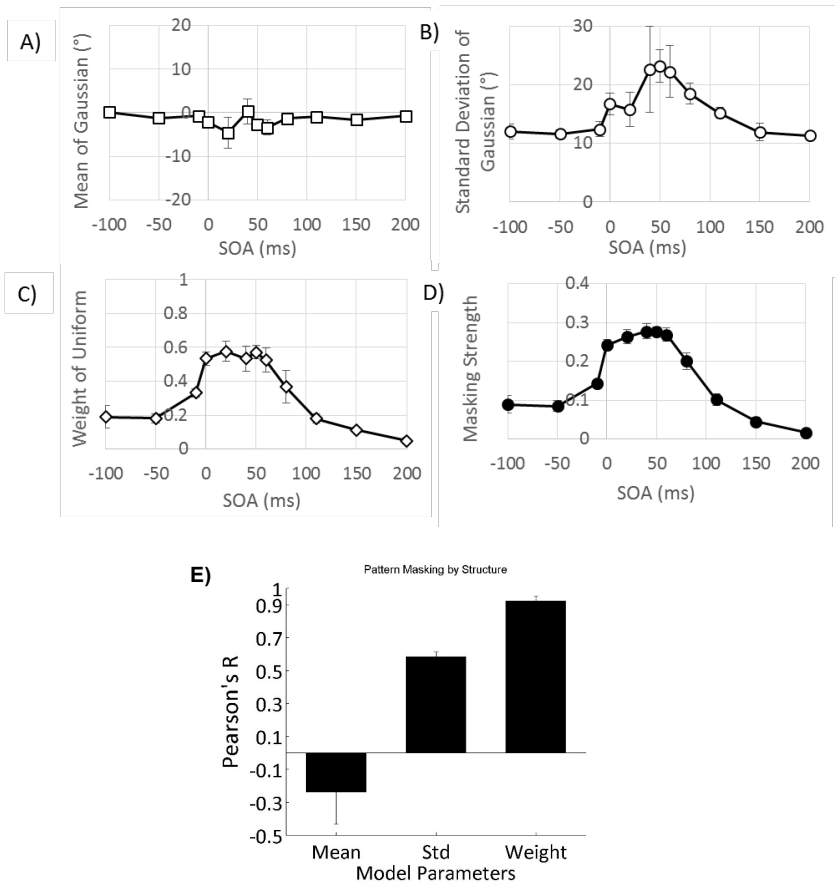

Figure A-6 Parameters of the GU model in pattern masking by structure. A) Means, B) Standard deviations of the Gaussian. C) Weight of the Uniform distribution. D) Masking strengths. E) Correlations between model parameters and masking strengths. Error bars represent ±SEM, n=5.

Appendix B

Table B-I shows the results of the BMC analyses for the experiments in Chapter 2. Individual BMC differences and corresponding Bayes factors are tabulated.

Table B-I BMC differences of all models from G model and corresponding Bayes factors in all masking types and for each observer are tabulated.

	Para-/Meta-contrast Masking		***Masking by Noise***		***Masking by Pattern***					
	ΔBMC	***Bayes Factor***	***ΔBMC***	***Bayes Factor***	***ΔBMC***			***Bayes Factor***		
Observer	*GU-G*	*GU/G*	*GU-G*	*GU/G*	*GU-G*	*GUCA-G*	*GUNN-G*	*GU/G*	*GUCA/G*	*GUNN/G*
SA	10.26	28684	1.39	4	9.02	6.16	5.54	8237	471	255
MNA	16.80	19681194	8.93	7523	14.19	11.13	10.63	1450658	68312	41168
FG	8.27	3909	4.88	132	8.08	5.26	4.55	3244	193	94
GJ	6.50	663	6.40	603	7.08	4.53	3.77	1190	93	43
MI	11.11	66838	10.24	27974	12.59	9.53	9.12	294355	13790	9096
Average	10.59	39633	6.37	583	10.19	7.32	6.72	26704	1514	827

Appendix C

We analyzed whether the parameter estimates of mixture models are reliable with the size of our data set in all experiments. First, as discussed in Chapter 2, two different techniques arrived at similar results. Moreover, in order to assess the reliability of our results with the given data size, we conducted the following simulation studies: First, we synthesized data from a single Gaussian (using the same number of data points as in the empirical data) whose standard deviation changes as a function of different conditions (to simulate different SOAs in the experiments). We then fitted the G and GU models to the synthetic data and we analyzed the BMC differences. We repeated this process (data generation and fitting) multiple times (N>100) and in all cases, model selection based on BMC differences successfully picked the G model as the winning model. An example is shown in Figure C-1. The leftmost panel shows the standard deviation of the Gaussian, that was used to generate the synthetic data (solid line) as well as the standard deviation of the Gaussian in the fitted GU model (markers). Right next to this, the weight of the uniform in the GU model is shown. There seems to be an overestimation of the weights at certain conditions (i.e., SOAs), however, this does not affect adversely the model selection process. In the third panel from left, actual and estimated standard deviations of the Gaussian in the G model are plotted. As expected they nicely match. Finally, in the rightmost panel, the BMC differences averaged across "synthetic subjects" (GU-G) are given. All points are below zero, indicating that the generated data is less likely to be drawn from a GU distribution.

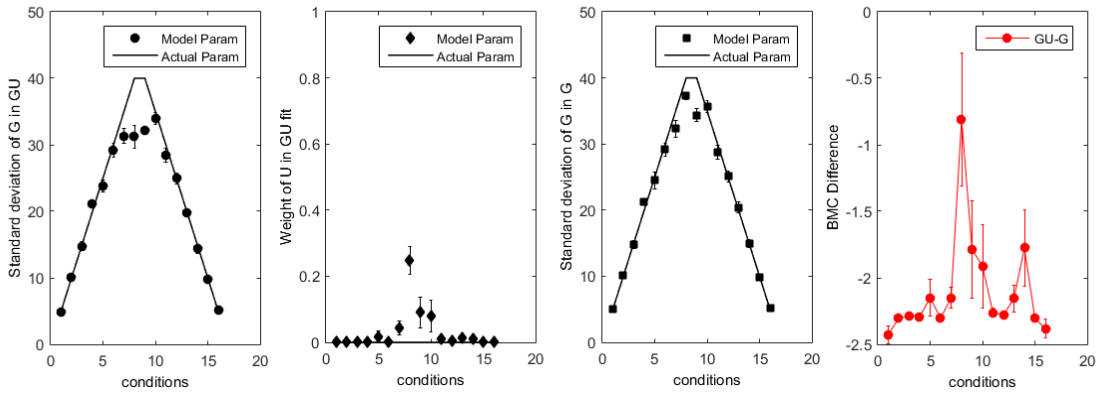

Figure C-1 Model fitting and model comparison by using synthetic data. See text for detailed explanations.

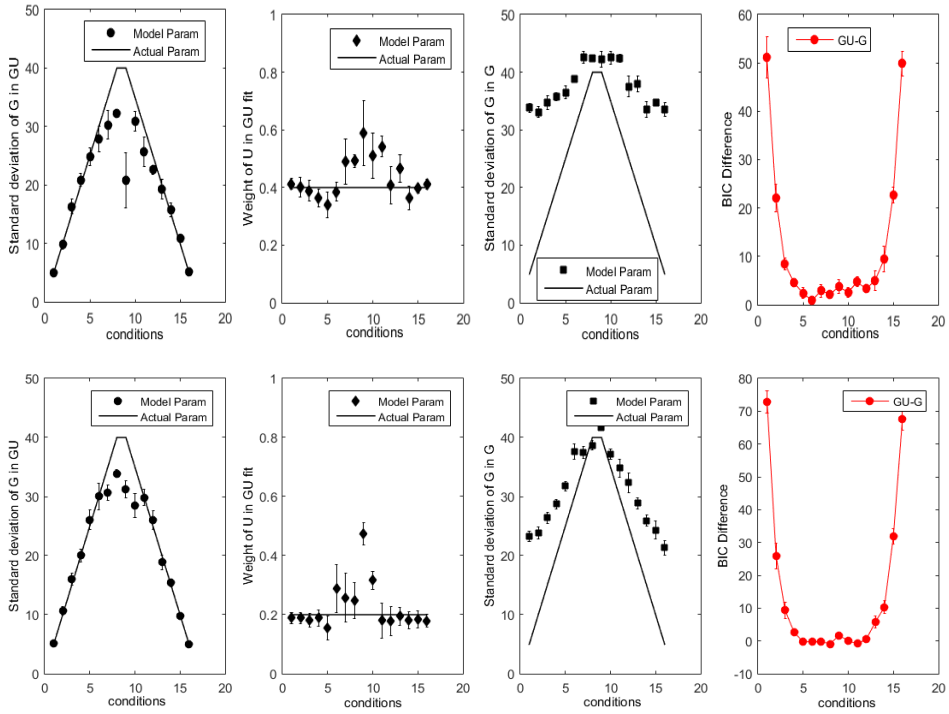

Figure C-2 Model fitting and model comparison by using synthetic data generated from a GU model with varying standard deviation for the Gaussian term as a function of SOA, and a constant weight for the Uniform component. See text for details.

Second, we synthesized data from a GU model with different weights of the Uniform component. We considered two different scenarios: (i) varying standard

deviation for the Gaussian term as a function of SOA, and a constant weight for the Uniform component (Figure C-2), (ii) a constant standard deviation for the Gaussian, and a varying weight for the Uniform (Figure C-3). We present two cases from each scenario to demonstrate that the usage of mixture models along with Bayesian model comparison techniques is warranted. The slight overestimation of the weight of the Uniform component at certain SOAs (depicted as "conditions" in the figures here) does not hinder selection of the correct source of the data at hand. In fact, whenever there is an overestimation, likelihood of the GU model dramatically decreases, as indicated by huge drops in BMC differences in the rightmost panels.

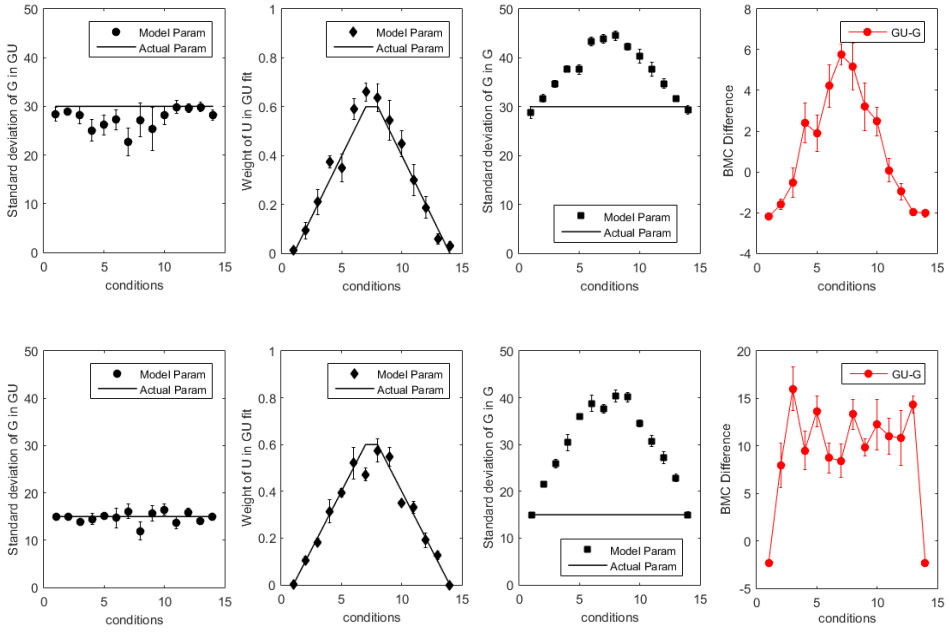

Figure C-3 Model fitting and model comparison by using synthetic data generated from a GU model with a constant standard deviation for the Gaussian, and a varying weight for the Uniform. See text for details.

Appendix D

Analysis of Baseline Data in the Set Size Experiment (Chapter 3)

Although we did not control the stimulus parameters to avoid ceiling effects in the baseline conditions, we performed a series of regression analyses on observers' performance in the baseline conditions using the models shown in Table 3-II. An ideal performance ceiling for the baseline conditions can be the performance with set size of one. However, we did not have this condition and hence, there might be ceiling effects for some of the observers, which could underestimate a potential effect of set size and SOA, and could give rise to spurious interaction effects. Table D-I summarizes the winning regression model for each observer. The fits of the winning models are shown by dashed lines in Figure 3-2 (the first column). In agreement with previous studies (Ogmen et al., 2013;Huyhn et al., in press), we found a main effect of set size for six of seven observers. In fact, the observer for which set size was not present in the best model (SOA and set size interaction alone was sufficient to fit data), was SA who is one of the authors and extremely trained on orientation judgment tasks. Therefore, failure to find an effect of set size for this observer might be due to the ceiling effect. Similarly, for five observers, a model with SOA term resulted in best fits. An effect of SOA in the baseline condition can be interpreted as (i) a memory effect, or (ii) a weak masking effect due to post-cue, or (iii) a combination of the two. Interestingly, for six observers, the best model included either or both of the interactions we investigated (i.e., SOA x set size or SOA^2 x set size). However, BIC differences between models with and without interactions terms were small (±2), indicating that interaction terms did not add much to the quality of fits.

See Figure D-1 for BIC differences between all regression models in the baseline condition.

Table D-I The best regression models for all observers in explaining the transformed performance.

	Best Regression Model	
Observer	**Baseline**	**Masking**
AK	M7	M16
CBK	M21	M16
EK	M9	M16
FG	M16	M16
GQ	M20	M16
MNA	M21	M16
SA	M4	M16

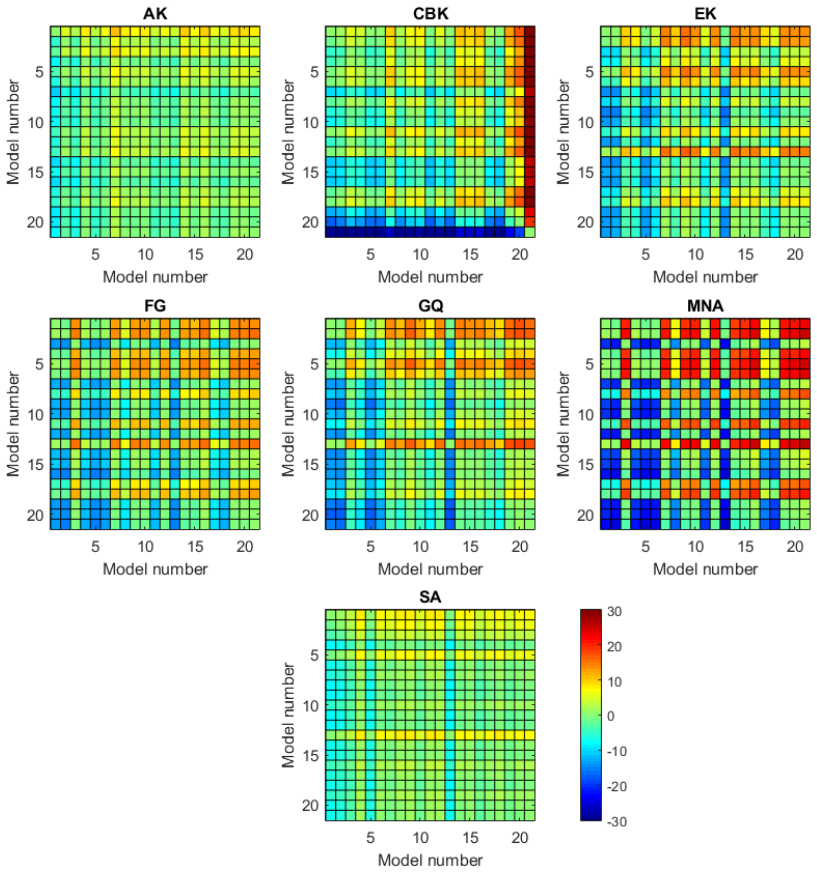

Figure D-1 BIC differences between pairs of regression models listed in Table 3-II. Each panel represents a different observer. All notations are the same as in Figure 4-2.

Appendix E

Here, we show the derivation of the BMC metric which was used to select the best performing model. Each model m_j produces a predicted error distribution $p(\boldsymbol{\varepsilon}|m_j, \boldsymbol{\theta})$, where $\boldsymbol{\varepsilon}$ is vector of observed response errors, and $\boldsymbol{\theta}$ is a vector of model parameters. For each model m_j, the likelihood of finding the observed errors averaged over model parameters is given by,

$$L(m_j) \triangleq p(\boldsymbol{\varepsilon}|m_j) = \int p(\boldsymbol{\varepsilon}|m_j, \boldsymbol{\theta}) p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}, \tag{E-1}$$

where $p(\boldsymbol{\theta}|m_j)$ represents the prior distributions for model parameters. Assuming that each response error is independent, (E-1) can be re-written as,

$$L(m_j) = \int \left[\prod_{i=1}^{N} p(\varepsilon_i|m_j, \boldsymbol{\theta})\right] p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}, \tag{E-2}$$

where N represents the number of trials, and ε_i represents the error in the ith trial. The maximum likelihood of a model within a parameter space is defined as,

$$L_{max}(m_j) = \max_{\boldsymbol{\theta}} L(m_j|\boldsymbol{\theta}). \tag{E-3}$$

To ensure that the integrand is of order 1 and thereby, to avoid numerical problems (Ester, Zilber, & Serences, 2015; Mackay, 2004; van den Berg, Shin, et al., 2012), we multiple and divide (E-2) by (E-3) to get (E-4) given below.

$$L(m_j) = \int \left[\prod_{i=1}^{N} p(\varepsilon_i|m_j, \boldsymbol{\theta})\right] \frac{L_{max}(m_j)}{L_{max}(m_j)} p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}. \tag{E-4}$$

It is convenient to take the logarithm of (E-4).

$$\ln L(m_j) = \ln\left(\int \left[\prod_{i=1}^{N} p(\varepsilon_i|m_j, \boldsymbol{\theta})\right] \frac{L_{max}(m_j)}{L_{max}(m_j)} p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}\right). \tag{E-5}$$

Since $x = \ln(e^x)$, (E-5) can be re-written as,

$$\ln L(m_j) = \ln\left(\int \exp\left(\ln\left\{\left[\prod_{i=1}^{N} p(\varepsilon_i|m_j, \boldsymbol{\theta})\right]\frac{L_{max}(m_j)}{L_{max}(m_j)}\right\}\right) p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}\right). \quad \text{(E-6)}$$

$$\ln L(m_j) = \ln\left(\int \exp\left[\ln L(m_j|\boldsymbol{\theta}) + \ln L_{max}(m_j) - \ln L_{max}(m_j)\right] p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}\right), \quad \text{(E-7)}$$

where $L(m_j|\boldsymbol{\theta}) = \prod_{i=1}^{N} p(\varepsilon_i|m_j, \boldsymbol{\theta})$. Since $e^{a-b} = e^a e^{-b}$, (E-7) can be re-written as,

$$\ln L(m_j) = \ln\left(\int \exp\left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j)\right] \exp\left[\ln L_{max}(m_j)\right] p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}\right). \quad \text{(E-8)}$$

Since $\exp\left[\ln L_{max}(m_j)\right] = L_{max}(m_j)$, and since $L_{max}(m_j)$ is independent of $\boldsymbol{\theta}$, it can be taken out of the integral.

$$\ln L(m_j) = \ln\left(L_{max}(m_j)\int \exp\left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j)\right] p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}\right). \quad \text{(E-9)}$$

Since ln (xy) = ln (x) + ln (y),

$$\ln L(m_j) = \ln\left(L_{max}(m_j)\right) + \ln\left(\int \exp\left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j)\right] p(\boldsymbol{\theta}|m_j) d\boldsymbol{\theta}\right). \quad \text{(E-10)}$$

Since we do not have an a priori reason to do otherwise, we used uniform priors for all model parameters. In mathematical terms, this can be expressed as follows.

$$p(\boldsymbol{\theta}|m_j) = \prod_{t=1}^{k} U(\theta_{t,min}, \theta_{t,max}), \quad \text{(E-11)}$$

where U(a,b) represents a uniform distribution over the interval [a,b], k represents the number of free parameters in the model m_j, and $\theta_{t,min}$ and $\theta_{t,max}$ represent the minimum and maximum boundaries for the t[th] free parameter. Since probability density function of a uniform random variable X ~U(a,b) is given by,

$$f(x) = \begin{cases} \frac{1}{R}, & x \le b \; and \; x \ge a \\ 0, & elsewhere \end{cases}, \tag{E-12}$$

where $R = (b - a)$, represents the range of X, (E-11) can be re-written as

$$p(\boldsymbol{\theta}|m_j) = \begin{cases} \prod_{t=1}^{k} \frac{1}{R_t}, & \cap_{t=1}^{k} [\theta_{t,min}, \theta_{t,max}] \\ 0, & elsewhere \end{cases}. \tag{E-13}$$

Substituting (E-13) to (E-10) yields,

$$\ln L(m_j) = \ln \left(L_{max}(m_j) \right) + \ln \left(\int_{Vol} \exp \left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j) \right] \left(\prod_{t=1}^{k} \frac{1}{R_t} \right) d\boldsymbol{\theta} \right), \tag{E-14}$$

$$\ln L(m_j) = \ln \left(L_{max}(m_j) \right) + \ln \left(\prod_{t=1}^{k} \frac{1}{R_t} \int_{Vol} \exp \left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j) \right] d\boldsymbol{\theta} \right), \tag{E-15}$$

$$\ln L(m_j) = \ln \left(L_{max}(m_j) \right) + \ln \left(\prod_{t=1}^{k} \frac{1}{R_t} \right) + \ln \left(\int_{Vol} \exp \left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j) \right] d\boldsymbol{\theta} \right), \tag{E-16}$$

and, finally,

$$\ln L(m_j) = \ln\left(L_{max}(m_j)\right) - \ln\left(\sum_{t=1}^{k} R_t\right) + \ln\left(\int_{Vol} \exp\left[\ln L(m_j|\boldsymbol{\theta}) - \ln L_{max}(m_j)\right] d\boldsymbol{\theta}\right), \tag{E-17}$$

where the integral is taken over a volume defined by the ranges of free parameters. $\ln L(m_j)$ given in (E-17) was taken as the BMC metric throughout this dissertation. For each statistical model m_j, we computed (E-17) and compared against each other.