MASKING IN CENTRAL VISUAL FIELD UNDER A VARIETY OF TEMPORAL AND SPATIAL CONFIGURATIONS

MARWAN DAAR

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

August 2017

ABSTRACT

For over a century, visual masking—where one stimulus reduces the visibility of another stimulus—has been used as a powerful tool to explore the visual system. Two major forms have emerged: backward masking and common onset masking. These two forms, which are characterized by the temporal properties of the stimuli, are often used to probe different underlying masking mechanisms, and the two forms typically employ a unique set of spatial characteristics of the mask. This clustering of stimulus properties makes it challenging to assess the effect of each stimulus property by itself. This dissertation describes an attempt to isolate the effects of these properties. In the first set of experiments various masking schedules are tested, including backward, common onset, and variations between, while keeping the spatial properties of the stimuli constant. In the second set of experiments four-dot common onset masking is explored in detail, and in one of the experiments, a single masking schedule is tested while varying the spatial properties of the mask. Across all experiments, target stimuli are presented foveally. A computational model is developed to account for data across both sets of experiments. Three important findings emerge. First, masking can be successfully obtained in central visual field using a variety of stimulus properties. Second, there is compelling evidence that persisting traces of these stimuli play an important role in masking. Third, there is strong evidence of both spatially local and global masking effects.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Hugh Wilson, for his patient mentoring over the many years I have worked in his lab. I will fondly remember the many hours I spent in your office, discussing data, learning from you as you gave an impromptu tutorial on your whiteboard, and chatting about your travels. I have always admired your deep understanding across a broad domain of knowledge, and your ability to come up with fascinating experiments. You are a wonderful role model as a scientist, and it has been a joy to get to know you on a more personal level over the years.

I would also like to extend my gratitude to my committee members, Dr. Frances Wilkinson, Dr. Richard Murray, Dr. Mazyar Fallah, and Dr. Jennifer Steeves. Your thoughtful feedback sharpened both my dissertation and mind. Dr. Vincent Di Lollo, it was an honour to have you on my examination committee. Your comprehensive feedback was extremely generous, and interacting with you was a joy that will stay with me.

David Nichols, Lisa Betts, Ingo Fründ, Krista Kelly, Minjung Kim, and Bob Hou, thank you for your friendship and mentorship. You have always been there when I needed to discuss an idea, learn a new concept, or enjoy the company of a friend. It has been wonderful sharing my grad school experience with you.

Teresa Manini, Lori Santos, and Freda Soltau, thank you for all you have done over the years. I would not have made it this far without you.

To my parents, Shahina and Abdallah, thank you for your support and gentle encouragement. Your presence has always been felt, unfaltering and loving.

To Olivia, Zoë, and Edie, the love and happiness we have shared has made life that much more beautiful, and I look forward to the adventure ahead!

Table of Contents

Abstract....ii

Acknowledgments....iii

Table of Contents....iv

List of Figures....vii

Chapter One....1

1.1 Masking, metacontrast, and masking functions....1

1.2 Integration and interruption....4

1.3 The sustained transient dual channel model....10

1.4 Object substitution masking....16

1.5 Object updating....20

1.6 Purpose of the current work....23

Chapter Two....27

2.1 Introduction....27

2.2 Experiment 2A....31

2.2.1 Stimuli....32

2.2.2 Procedure....34

2.2.3 Results....36

2.3 Experiment 2B....37

2.3.1 Results....39

2.3.2 Discussion....40

2.4 Experiment 2C....40
2.4.1 Procedure....42
2.4.2 Results....42
2.5 Experiment 2D....44
2.6 Experiment 2E....46
2.7 Computational Model....48
2.8 Experiment 2F....54
2.8.1 Results....56
2.8.2 Discussion....57
2.9 General Discussion....58

Chapter Three....67
3.1 Introduction....67
3.2 Experiment 3A....69
3.2.1 Stimuli....70
3.2.2 Procedure....72
3.2.3 Results & Discussion....73
3.3 Experiment 3B....74
3.3.1 Procedure....74
3.3.2 Results & Discussion....76
3.4 Experiment 3C....77
3.4.1 Procedure....78
3.4.2 Results & Discussion....79
3.4.3 Mixture Model Analysis....83
3.5 General Discussion....86

Chapter Four....93

4.1 Model Description....93

4.2 Model Performance....102

4.3 Discussion....107

Chapter Five....112

5.1 Summary of findings....112

5.2 Limitations of computational model....117

5.3 Does the model implement mask blocking?....118

5.4 Future directions and final thoughts....122

References....124

List of Figures

Figure 1.1. Schematic depiction of McDougall's (1904) apparatus.....3

Figure 1.2. Two classes of masking function: type A and B, are shown here.....4

Figure 1.3. Three different integrative masking mechanisms are shown here.....6

Figure 1.4. Data from Crawford's (1947) masking by light experiments.....15

Figure 1.5. Common onset masking stimuli from Di Lollo, Bischof, & Dixon (1993)….17

Figure 2.1. Two broad classes of masking schedules.....29

Figure 2.2. Target and mask stimuli.....33

Figure 2.3. General schematic of backward and trailing mask conditions in Experiment 2A.....34

Figure 2.4. Trial sequence for backward and trailing masking conditions in Experiment 2A.....36

Figure 2.5. Data from four observers in Experiment 2A.....37

Figure 2.6. Schematic for conditions in Experiment 2B.....38

Figure 2.7. Data from Experiment 2B.....39

Figure 2.8. Conditions for Experiment 2C.....41

Figure 2.9. Data from Experiment 2C....43

Figure 2.10. Conditions for Experiment 2D....45

Figure 2.11. Data from Experiment 2D....45

Figure 2.12. Three conditions from experiment 2E are shown....48

Figure 2.13. Top Left: Energy of mask stimulus (solid line) overlaid with modeled magnocellular response to this stimulus (dotted line) for the Gap condition....53

Figure 2.14. Transition from type B to type A backward masking....56

Figure 2.15. Conditions for Experiment 2F. In the Pulse conditions, the mask briefly doubles in contrast relative to the rest of the mask trail....59

Figure 2.16. Data from Experiment 2F....61

Figure 3.1. Target and four-dot mask drawn to scale....75

Figure 3.2. Trial sequence and results for Experiment 3A....79

Figure 3.3. Stimuli used in Experiment 3B....81

Figure 3.4. Results from Experiment 3B....84

Figure 3.5. Orientation bias results from Experiment 3C....87

Figure 3.6. A: Distribution of errors across nine observers for the common offset and 250 ms conditions, along with model fits....92

Figure 4.1. Magnocellular and Parvocellular receptive fields.....109

Figure 4.2. Magno and parvocellular temporal impulse response functions.....112

Figure 4.3. Magno and parvocellular response functions (N(t)), along with stimulus timecourses shown in green (T(t)).....112

Figure 4.4. Contrast gain functions for the magno and parvocellular components of the model.....114

Figure 4.5. A visualization of the transformation from the original face stimulus into the magnocellular and parvocellular components.....115

Figure 4.6. Backward and common onset masking functions using a disc and annulus, shown at top.....118

Figure 4.7. Model output vs. face masking data from Experiment 2A.....119

Figure 4.8. Model output vs. face masking data from Experiments 2C & 2D.....120

Figure 4.9. Model output vs. Experiment 3B.....121

Figure 4.10. Left: Backward masking simulations as a function of mask duration.....125

Figure 5.1. Comparison of output of computational model (Chapter 4) with data from Francis & Cho (2007).....136

Chapter One

Introduction

1.1 Masking, metacontrast, and masking functions

If you were to glance at any point in your environment, there would likely be objects that would, were it not for the presence of an occluding object, be readily visible to you. Here, we can account for the reduced visibility of the occluded object by understanding the way light interacts with matter: the more opaque the occluding object is to visible light, the less of that light will reach the eyes of the observer. Often, however, we must turn to the observer's visual system itself in order to understand how the visibility of some features in the world can be modulated by others. For example, in order to explain a phenomenon such as simultaneous brightness induction, where the luminance in one area of the visual field can alter the brightness of another area (Heinemann, 1955), it is helpful to understand something about the topology of visual neural networks, such as the structure of receptive fields (Blakeslee, Pasieka, & McCourt, 2005), along with relevant biophysical properties, such as those involved in neural inhibition (Jonas & Buzsaki, 2007). Visual masking refers to the class of phenomena wherein the visibility of one visual stimulus (the target) is impaired by the presence of another stimulus (the mask). The study of visual masking thus explores masking phenomena and their regularities, as well as the underlying properties of the visual system. In the following introductory sections, I will present a historical overview of this field, emphasizing the relationship between the employed stimulus properties and

the ensuing theoretical developments. At the end of the introduction, I will present the motivation for my own experiments.

Around the turn of the 19th century, researchers discovered that visual stimuli could influence the perception of other stimuli not only across space (e.g. simultaneous contrast), or time (e.g. successive contrast), but across *both* space and time. One approach to studying such stimuli was by use of a rotating disc, which contained apertures through which an illuminated back surface could stimulate the eyes of an observer (McDougall, 1904). By varying the radial distance between these apertures, one could control the spatial separation between the stimuli, while the temporal separation was modulated by varying the arc length between them, along with the speed of disc rotation (Figure 1.1). It had already been observed that stimuli left a persisting visual impression, even after the physical stimulus disappeared (Exner, 1868, as cited in Breitmeyer & Öğmen, 2006). In the context of McDougall's apparatus, this meant that with a single aperture, a dull band that lagged the aperture was visible for a brief period. This "afterimage", or *metaphotic* component was distinguished from the primary image, or *homophotic* component. When two apertures were present, which were appropriately separated in time and space, the afterimage of the leading stimulus could be suppressed by the onset of the trailing stimulus. To distinguish this from the idea of simultaneous contrast, the term metaphotic contrast, or *metacontrast* was used to describe this homophotic on metaphotic suppression (Stigler, 1910, as cited in Breitmeyer & Öğmen, 2006). While our understanding of the visual system has evolved considerably since this period, these ideas did anticipate some more modern theories of masking, for example Weisstein's two factor model (Weisstein, Ozog, & Szoc, 1975), and Breitmeyer and Ganz's dual channel model (Breitmeyer & Ganz, 1976).

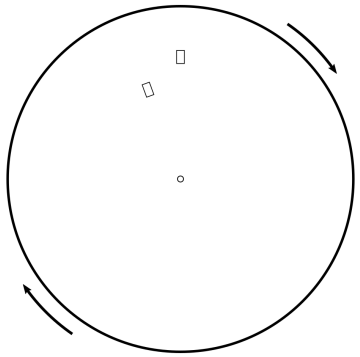

Figure 1.1. Schematic depiction of McDougall's (1904) apparatus. Two apertures are shown here, through which an illuminated back surface projects light. As this back surface (not shown) is comparable in area to that of the rotating disc, the apertures are constantly transilluminated. To serve as a fixation point, a porcelain bead was hung from a thread and illuminated by a separate light source. This fixation point is depicted by a small circle here.

Advancements in display technologies, such as electronically programmable tachistoscopes and cathode ray tubes enabled researchers to efficiently explore a wide range of stimulus configurations with high spatial and temporal precision, yielding rich insights into masking and the visual system. At the heart of many of these investigations is the masking function, which describes some metric related to performance, such as brightness judgments or contrast detection thresholds, as a function of some property of the stimuli—often the stimulus onset asynchrony (SOA) between the target and mask (Figure 1.2). One of the most intriguing aspects of masking is that these functions often display a masking peak when the onset of the target precedes that of the mask by a particular interval. Such a function is known as a type B backward masking function (Kolers, 1962). Furthermore, the shape of the masking function may be affected by the type of mask used. For example, a noise mask tends to produce a peak of masking at an SOA of 0 (e.g. Agaoglu, Agaoglu, Breitmeyer,

& Öğmen, 2015), known as type A masking. The shape of these masking functions, and the way they can be influenced by varying the stimulus properties, are important clues that continue to spur theoretical development within the field of vision and masking.

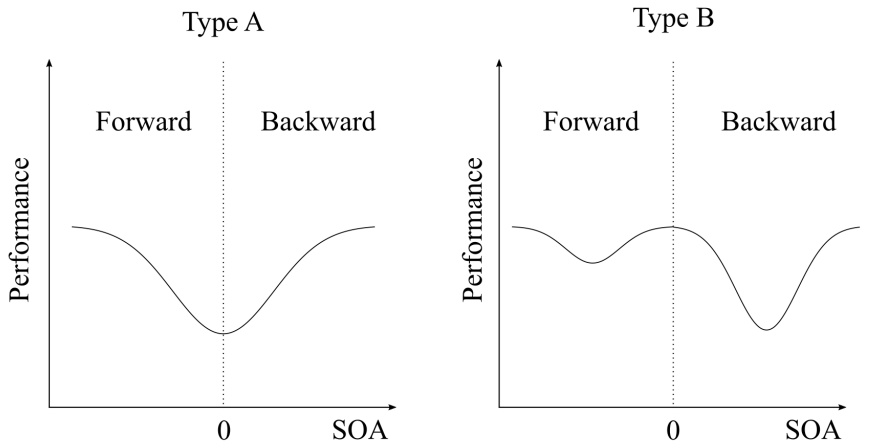

Figure 1.2. Two classes of masking function: type A and B, are shown here, as a function of stimulus onset asynchrony (SOA). Performance is inversely related to masking strength. Forward masking refers to SOAs at which the mask appears before the target, and backward masking occurs when the target appears before the mask. Figure adapated from Breitmeyer & Öğmen (2006).

1.2 Integration and interruption

Theoretical accounts of masking were historically couched in terms of *integration* and *interruption* (Scheerer & Bongartz, 1973). In integration, the target and mask are perceived as a single composite, due to the limited temporal resolution of the visual system. While in some cases this can improve the visibility of the target (e.g.

subthreshold summation), in other cases visibility is reduced (Figure 1.3). A simple case of this is contrast reduction due to luminance summation (Eriksen & Hoffman, 1963). For example, in masking by light, where a target presented on a uniform background is followed by a brief uniform field of light, the contrast of the target against the background is reduced when the two presentations are integrated into a composite image. Target clarity can also be compromised when the target plus mask composite makes it difficult to clearly delineate which contours uniquely belong to the target. Here, the mask is essentially camouflaging the target. Finally, when a non uniform mask spatially overlaps a target, the target is often degraded, as for example with a noise mask. It is important to recognize that integration per se is not a mechanism of masking, but rather sets the stage wherein any number of masking mechanisms can play their role. According to this account of integration, the closer together in time the target and mask are, the more likely they are to be perceived as a single composite and/or the stronger the composite will be. As such, any mechanisms that depend on integration will exert their effects maximally when the target and mask are simultaneous, and will therefore contribute towards type A masking functions.

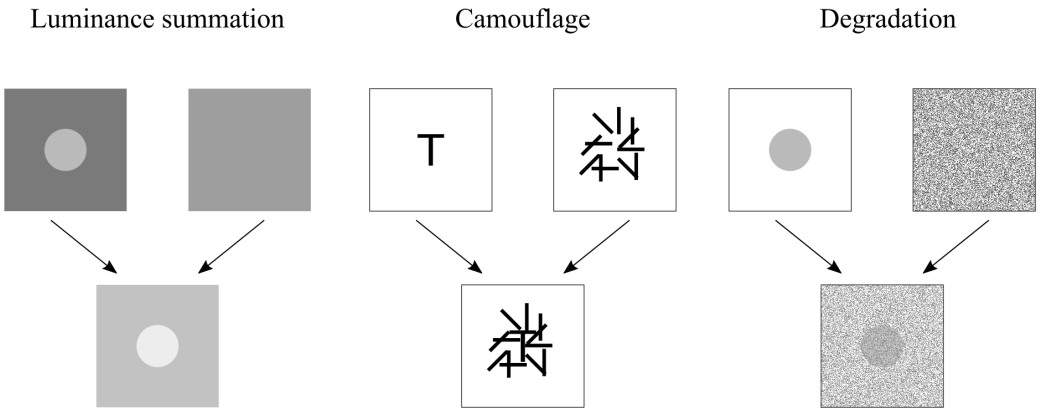

Figure 1.3. Three different integrative masking mechanisms are shown here. In each case, the target and mask fields are shown on the left and right, respectively. The combined field is shown below each pair. In the luminance summation shown here, the Michelson contrast of the target disc against the background (before being integrated with the masking field of light) is 0.43. After integration, it is reduced to 0.22 (these values assume the figure is rendered with a gamma of 2.2). Note that these values assume that the two fields are combined simultaneously (i.e. superimposed), rather than presented successively and integrated across time.

Interruption, unlike integration, is itself considered a mechanism of masking. The general idea behind interruption is that the mask interrupts target processing before it can be consolidated into awareness (Averbach & Sperling, 1961; Sperling, 1963). If such interruption occurs before enough information has been extracted in order that the observer can consciously report the target's content, then masking is said to have occurred. Importantly, due to persistence, information can continue to be extracted even after the target disappears, and this means that the target is vulnerable to masking even if the mask is presented after target offset. In Sperling's (1963) work, where the target comprised an array of letters, the number of letters observers were able to recall

increased linearly with the delay between the target and and a noise mask, with peak masking occurring at simultaneous presentation. Sperling's interpretation of the data was that the letters were scanned into an accessible memory store at a certain rate (e.g. 10 ms per letter), and that when the noise mask appeared, no more letters could be consolidated.

While Sperling (1963) observed peak noise masking at simultaneous presentation of target and mask, type B effects observed in the early 20th century were interpreted as being due to latency differences between a high energy mask and a relatively low energy target (Piéron, 1935, as cited in Kinsbourne & Warrington, 1962). Here, the mask, presented to the visual system after the target, catches up, or "overtakes" the target due to the mask's lower latency, which in turn is due to the mask's higher energy (Monnier, 1952). According to these ideas, the time window, or critical interval, over which a mask can effectively interrupt a target varies with the relative energies of the target and mask. Kinsbourne & Warrington investigated this, using noise masks and letter targets, and found that this critical interval was best modeled as being inversely proportional to the target energy (which was controlled by target duration), and relatively unaffected by mask energy. Thus, the critical interval multiplied by target energy was a constant. This finding, rather than supporting a theory of interruption occurring over different relative latencies, is compatible with a (temporal) integration account of masking, where successful perception depends upon temporally integrating sufficient target energy. Here, longer duration targets, which can be integrated into a stronger percept compared to a shorter duration target, achieve immunity from masking relatively early on after target offset. Turvey (1973) explored this further in a series of experiments, and in his second experiment, confirmed the prediction that the energy x critical interval law held even when target energy was modulated via luminance, rather

than duration, thus strengthening the integration account. Importantly, however, in conditions that isolated target mask interactions to non peripheral loci (e.g. dichoptic presentation), he found a different relationship held. Here, target duration + critical interval was a constant. This second relationship, which is essentially identical to the findings of Kahneman (1967)[1], and which is known as the SOA law, will be important when we discuss Breitmeyer and Ganz's (1976) dual channel model (Chapter 1.3).

Direct evidence for integration masking came from Schultz & Eriksen (1977), who used dotted stimuli to compare the effect of different types of noise masks on target identification. The targets were partially degraded to begin with (due to missing dots) and the noise masks were designed such that a composite image of the mask and target would either enhance or disrupt the target, by either filling in the missing gaps or adding non-informative clutter. The results showed that with the informative mask, target identification was easiest at simultaneous presentation, and became more difficult as the delay between target and mask increased. With the non-informative mask, the opposite was true: performance was worst at simultaneous presentation, and improved as a function of delay. Furthermore, this pattern of results held regardless of whether the mask followed the target (backward masking), or whether the target followed the mask (forward masking). These results suggest that performance was predicated on the observer integrating the target and mask into a composite image: the longer the two stimuli were separated in time, the less effective this integration (and thus the less enhancement or disruption occurred, depending on the mask type). In another study, however, Navon & Purcell (1981) found strong evidence that masking could occur in the *absence* of integration. Their masking stimulus consisted of a pattern of lines, a subset

1 Kahneman (1967) was interested in the relationship between apparent motion and metacontrast, and so studied the SOA law as it related to both phenomena.

of which was completely superimposed on a target letter. Importantly, in a composite image of the target and mask, target contrast was not reduced, and was in fact enhanced. In a condition where the mask and target were different hues (red and blue, respectively), they found evidence of integration at an SOA of 10 ms, as the spatially coincident regions were perceived as a single purple hue, but at not at an SOA of 50 ms, where the red and blue stimuli were registered independently. Yet target identification performance at the smaller SOA was better than at the larger one. The authors proposed a model whereby interruptive mechanisms, that become weaker as a function of SOA, are protected against by integration at smaller SOAs. When integration itself does not lead to masking (e.g. through contrast reduction, camouflage, etc.), the combination of this "fortunate" integration, and disruptive interference can result in type B masking curves: at smaller SOAs, where the two stimuli are processed as a composite, interruption does not occur; at larger SOAs, where the stimuli are individuated by the visual system, the leading one is interrupted by the trailing one.[2] While Navon and Purcell's model has limitations (see Breitmeyer & Öğmen, 2006), this story of a tension between integration and individuation has recently emerged in the development of the theory of object substitution masking (Goodhew, 2017; Goodhew, Pratt, Dux, & Ferber, 2013), although as we shall later see, the consequences for successful target perception are rather different in this case (Chapter 1.5).

The question of whether a mask operates through interruption or integration is an important one (Eriksen, 1980). If a researcher wishes to use such a mask to limit target processing time (e.g. Reicher, 1969), then this would only make sense if interruption plays a significant role. Thus, a fair amount of research (e.g. Liss, 1968;

2 This same explanation was given twenty years prior, to account for the effect an annulus mask had upon a partial report task (Averbach & Sperling, 1961).

Schultz & Eriksen, 1977; Spencer & Shuntich, 1970) was dedicated to uncovering whether masking operated through integration, interruption, or both. However, much of this research tended to be limited with respect to the variety of mask under study, and it is likely that the theoretical development of masking was influenced by the types of masks that were used. A further issue was a lack of conceptual clarity surrounding the notions of integration and interruption. For example, while integration theories often postulate that peak masking occurs at simultaneous presentation (i.e. when the target and mask are most "fully" integrated), broader accounts of integration, where, for example, signals with different timecourses sum together to weaken target perception (e.g. see Kahneman's, 1968 characterization of Weisstein, 1968), can account for type B masking. Yet, the presence of type A masking, particularly type A masking that is symmetric with respect to backwards and forwards masking, has been used as a *test* for integration masking (Navon & Purcell, 1981). Moreover, interruption has often been conceptualized merely as a mechanism that is separate from integration. Here, if there is evidence of masking that occurs outside of integration, it is simply assumed that interruption has occurred. While some researchers are forthcoming about this negative definition of interruption (Navon & Purcell, 1981), others seem to have implicitly assumed a particular mechanism of interruption simply by the presence of masking in the absence of integration (Spencer & Shuntich, 1970).

1.3 The sustained transient dual channel model

A number of models of masking have been proposed over the last half century (Bridgeman, 1978; Kahneman, 1967; Weisstein et al., 1975), and some of them have brought clarity to the aforementioned issues. For a review of these and other models, see Breitmeyer & Öğmen (2006). One of the most successful of these—the sustained

transient dual channel model (Breitmeyer, 1992; Breitmeyer & Ganz, 1976; Breitmeyer & Öğmen, 2006)—provides an impressive amount of explanatory power, and makes use of the fact that there are two distinct channels in the visual system that process information in parallel, a feature which lies at the heart of the model. The magnocellular (transient) pathway receives input from the parasol ganglion cells in the retina. As these cells pool input from a large number of photoreceptors, information that is carried through this pathway has a relatively limited spatial resolution. The phasic response of these cells, combined with fast axonal conduction velocities that characterize the projections to the lateral geniculate nucleus (LGN) reflect an ability to rapidly convey information with high temporal resolution. In contrast, the parvocellular (sustained) pathway, which receives input primarily from midget ganglion cells, is characterized by sensitivity to fine spatial detail, and has a relatively high latency, tonic response. The contrast gain characteristics of these pathways differ markedly (Kaplan & Shapley, 1986): magnocellular ganglion cells have a low contrast threshold, and show a high gain and compressive nonlinearity that saturates at relatively low contrasts. Parvocellular cells have a higher contrast threshold, a shallower gain, and a more linear response until saturation at a relatively high contrast. Magnocellular (M) and parvocellular (P) cells project to the dorsal and ventral pathways, respectively, although this separation is not exclusive (Ferrera, Nealey, & Maunsell, 1992; Merigan & Maunsell, 1993), and there is a degree of convergence across their inputs (Vidyasagar, Kulikowski, Lipnicki, & Dreher, 2002). The combined characteristics of these pathways make the magnocellular pathway especially suited to detecting sudden changes in luminance (such as the appearance of an object), motion, and location, while the parvocellular pathway excels at processing form, and discrimination of color and luminance. Masking, according to the dual channel model, results from the interaction

within and between these channels, and occurs primarily in the following ways: transient on sustained (interchannel) inhibition, via lateral inhibitory connections; sustained on sustained (intrachannel) inhibition, via surround on center antagonism; and the sharing of sustained or transient pathways when mask and target are overlapping (e.g. integrative mechanisms).

Together, these mechanisms can account for a wide range of masking phenomena. For example, type B backward masking can be explained by the high latency sustained response of a target being temporally superimposed, and thus optimally inhibited by, the low latency transient response of a subsequently flashed mask. This account of masking, which is a version of the overtake hypothesis, is consistent with the SOA law described earlier (Turvey, 1973). Type B *forward* masking[3] (also known as paracontrast, see Figure 1.2) can be accounted for by the finding that the surrounds of classical receptive fields of parvocellular retinal ganglion cells lag the response of their centers by around 10-30 ms (Benardete & Kaplan, 1997; Maffei, Cervetto, & Fiorentini, 1970; Singer & Creutzfeldt, 1970). The finding, however, that paracontrast effects are observed dichoptically (Kolers & Rosner, 1960) and at SOA magnitudes of between 200 ms and 450 ms (Breitmeyer et al., 2006; Öğmen, Breitmeyer, & Melvin, 2003) demonstrates that the locus of any such inhibitory interactions must also occur beyond the retina, and are likely subserved by slower cortical inhibitory interactions (Connors, Malenka, & Silva, 1988).

As the shape of a masking function reflects the underlying masking mechanisms, it is possible to alter a masking function by biasing different mechanisms. One way to do this is by varying the types and properties of the mask. For example, masks that

3 Unless specified otherwise, the use of type A and B terminology in this dissertation refers to backward masking.

reduce target visibility when integrated with the target (e.g. masking with a flash of light, or with noise), tend to produce type A functions, whereas masks whose contours are closely aligned with those of the target can produce type B functions. In this latter case, transient on sustained inhibitory mechanisms are facilitated by the close proximity between the mask and target contours, especially when these contours are parallel (Ganz, 1966). The dual channel model elegantly accounts for the observed shift from type B to type A functions that occurs when the ratio of mask to target energy is increased (Breitmeyer, 1978; Growney & Weisstein, 1972; Stewart & Purcell, 1974). As the response of M cells saturates at lower contrasts than that of P cells (Kaplan & Shapley, 1986), increasing mask contrast can result in a relative shift in favour of those mechanisms that depend on the strength of the parvocellular response (e.g. luminance summation, center-surround antagonism). As these mechanisms peak at smaller SOAs than interchannel inhibition, the peak of the masking function will shift backwards accordingly, when mask contrast is increased.

Masking functions also reflect the spatial frequency properties of the stimuli. There is an abundance of electrophysiological and psychophysical evidence suggesting that the previously discussed cluster of properties that distinguish the two channels is joined by a differential response to spatial frequency, where the M and P pathways respond to low and high spatial frequencies, respectively (Ikeda & Wright, 1975; Legge, 1978; Merigan, Katz, & Maunsell, 1991; Wilson, McFarlane, & Phillips, 1983; although see Skottun, 2015). Rogowitz's (1983) masking data conform well to predictions based on these properties. She found that as the spatial frequency of a target increased, peak masking tended to occur at higher SOAs. This makes sense given that two properties that are shared by the P pathway are a high response latency, and a preference for higher spatial frequencies. A high spatial frequency target, which biases target

processing in favour of the P pathway, will therefore be optimally inhibited by a mask that follows with a longer delay than would be required with a low spatial frequency target. In other words, the use of a high spatial frequency target essentially slowed down target processing, and this was reflected by peak masking at a greater SOA. Rogowitz also found that as the spatial frequency of the *mask* increased (thus favouring mask processing along the P pathway), peak masking shifted to smaller SOAs and became weaker. In other words, a higher spatial frequency mask slowed down mask processing. In another study, Green (1981) found that when a grating was masked by a 700 ms pulse of light, performance thresholds as a function of SOA showed overshoots at the onset and offset of the light mask when the grating was of a low, but not high spatial frequency. This suggests that the magnocellular activity, generated by the dramatic luminance increment and decrement, interfered with magnocellular mechanisms responsible for processing the low spatial frequency grating, perhaps by adding noise to this channel (Breitmeyer & Öğmen, 2006, pp. 188-9), but did not interfere with those mechanisms tuned to the high spatial frequency grating. This finding, and in particular the shape of the masking functions, is a beautiful example of how masking can provide a keen window into the visual system, and is reminiscent of Crawford's (1947) data (Figure 1.4).

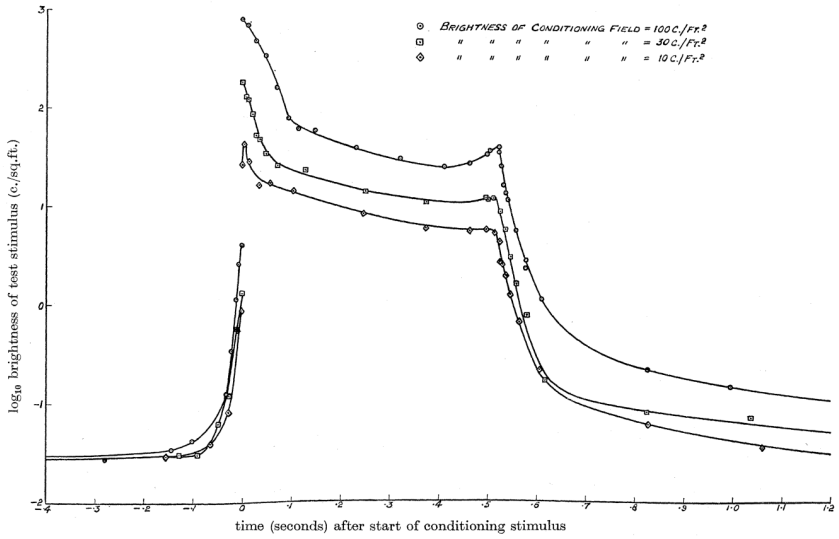

Figure 1.4. Data from Crawford's (1947) masking by light experiments. Plotted here are brightness thresholds of a 10 ms test flash as a function of the time between the target and mask onsets. The mask (conditioning field) was flashed for a duration of 524 ms. Note that positive values on the abscissa mean that the target was flashed after the mask. The transient overshoots, or "ears", likely reflect magnocellular interactions between the target and mask. The large number of SOAs tested here offers an impressive view of the masking phenomenon. Image reproduced from Crawford (1947).

The nature of any particular task has an impact on which perceptual mechanisms are employed, and changing task requirements can therefore change the shape of a masking function. Fehrer & Raab (1962) found that simple reaction times to a target remained unchanged even at SOAs that produced profound changes in target brightness (see also Fehrer & Biederman, 1962; Harrison & Fox, 1966). Here, the type B function found with brightness suppression can be explained by transient mask activity inhibiting sustained target activity. The transient target activity responsible for target

detection, however, was not inhibited at SOAs that produced peak brightness suppression. In a more recent experiment, Breitmeyer et al. (2006) found that both brightness judgments and contour discriminations of a target showed type B masking functions, but that peak masking was about 25 ms earlier in the contour discrimination task, a finding that comports with the idea that boundaries, or contours, are processed in advance of surfaces (Breitmeyer & Tapia, 2011; Breitmeyer, 2014; Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999).

1.4 Object substitution masking

One of the first significant challenges to the sustained transient model was the finding that a foveally presented target could be strongly masked by a common onset mask with a delayed offset (Bischof & Di Lollo, 1995; Di Lollo, Bischof, & Dixon, 1993). That is, the target and mask appeared together, and the mask persisted for a variable duration after target offset, with longer durations producing greater masking (Figure 1.5). This finding is hard to explain away with integration (where a longer duration mask is equivalent to increasing mask contrast, e.g. Breitmeyer, 1978), as the masks were brightness matched across the different durations. Moreover, as the masks comprised contours that were parallel and nonoverlapping with the target (a contour mask), and the task was to report which side of the target had a gap, it is unlikely that an integrated composite of these stimuli would have radically impeded performance.

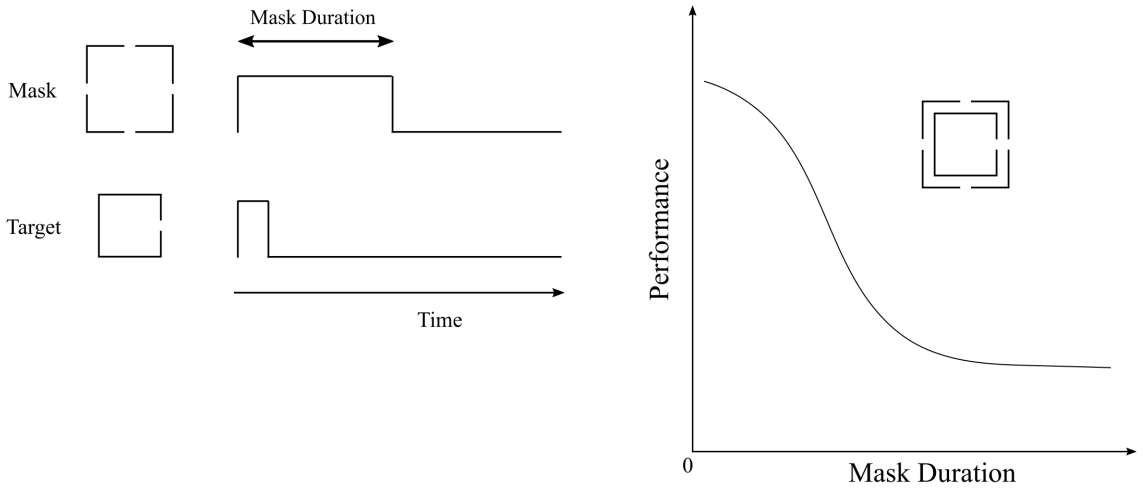

Figure 1.5. Common onset masking stimuli from Di Lollo, Bischof, & Dixon (1993); Bischof & Di Lollo (1995). On the **left** are shown the target and mask stimuli, along with their timecourses. On the **right** is shown a typical common onset masking function.

This brought into serious doubt the claim that SOA is the key variable that dictates masking with a contour mask, and cannot be readily explained by the sustained transient model (although, as Bischof & Di Lollo (1995) showed, Bridgeman's (1978) model of recurrent lateral inhibition is able to). Follow up work revealed that a sparse, four-dot mask, presented parafoveally and in unpredictable locations, could produce powerful backward masking (Enns & Di Lollo, 1997); and that such a mask could mask a target under common onset conditions (Di Lollo, Enns, & Rensink, 2000). Furthermore, these latter two studies found that masking showed little to no sensitivity to the spatial separation between target and a four-dot mask, suggesting that spatially sensitive mechanisms, such as those involved in lateral inhibition between mask and target contours, are not behind the observed masking effect.

To explain these findings, Enns & Di Lollo (1997) described masking in terms of "object substitution", an interruptive mechanism whereby the mask replaces the target as the focus of conscious processing. This idea was elaborated further in Di Lollo et al. (2000) where object substitution was conceptualized within a reentrant processing framework (Mumford, 1991, 1992). Here, successful perception depends upon interactions between feedback from higher visual (and frontal) cortical areas, and information in lower visual areas. For example, when a target and mask are presented simultaneously, a rapid and coarse representation of the image is rapidly shuttled to higher brain regions where a set of perceptual hypotheses are generated and sent back down to lower visual areas. In the case where the target and mask offset together, these signals will be met with a decaying representation of the target mask composite, and will likely match one of the hypotheses. As such, perception of the target mask composite will result. If, however, the mask remains visible after target offset, the incoming feedforward information (mask alone) will conflict with any hypotheses that include the target. Here, this mismatch will result in a further set of feedforward feedback iterations, favouring conscious perception of the mask alone. Based on this idea, Di Lollo et al. were able to successfully model their data with their computational model of object substitution (CMOS). In this model, the set size of the stimulus array influences the attentional resources available for target perception: higher set sizes means that attention is more sparsely distributed, and thus more time (more iterative loops) is required to consolidate a percept. With this reduced attention, there is a higher probability that any perceptual hypotheses that are fed back to lower visual areas will encounter the mask alone before successful target perception has occurred. Their data and model both confirmed the prediction that there would be an interaction between set size and the magnitude of masking, where larger set sizes produce greater masking, and

where masking is measured as the difference in performance between a common offset and a delayed offset mask.

While the findings that led to object substitution were unique—to my knowledge there is no published data before Di Lollo et al. (1993), for example, where common onset masking has been measured as a function of brightness matched duration—they do not by themselves demonstrate reentrant processing, at least not in the way that CMOS proposes. Francis & Hermens (2002) showed how a number of other models can account for the main finding in Di Lollo et al. (2000)—that a common onset sparse mask produces masking that increases as a function of mask duration, and that set size interacts with this effect (although it is notable that in a later study, Francis & Cho (2007) found that these same models failed to predict the pattern of results found with *backward* masking with a sparse mask, whereas CMOS succeeded). While some, but not all (e.g. Weisstein, 1972) of these models feature some form of reentrant processing, the function of such processing in these other models is rather different from the hypothesis testing role that CMOS purports. Object substitution theory, however, is certainly consistent with evidence showing pervasive feedback projections in the brain and visual system (Felleman & Van Essen, 1991). Furthermore, there is growing evidence that reentrant processing can not only modulate activity in lower visual areas (Ringach, Hawken, & Shapley, 1997), but that it is actually a hallmark of conscious perception (Lamme, 2006; Lamme & Roelfsema, 2000). Of particular interest is the finding that prefrontal brain areas have shown activation in advance of activity in the temporal cortex, and that this pattern of activity is associated with successful recognition of a masked object (Bar et al., 2006). Similarly Fahrenfort, Scholte, & Lamme (2007) found that late stage posterior occipital activation (>100 ms after target onset), which was

present on unmasked trials, and which presumably reflected reentrant processing, was abolished when the target was masked.

1.5 Object updating

Object substitution masking (OSM) originally denoted a particular set of masking mechanisms. Since the original findings (Di Lollo et al., 2000; Enns & Di Lollo, 1997), however, the term has often been associated with the spatial and temporal parameters of the stimuli that are employed, rather than the purported underlying mechanisms. In other words, if a common onset delayed offset four-dot mask produces masking, this is often termed OSM even if mechanisms other than the ones originally proposed are thought to be responsible (e.g. see Goodhew, Edwards, Boal, & Bell, 2015). In the two decades since Enns & Di Lollo's original paper, object substitution theory has undergone significant developments (for reviews, see Goodhew, 2017; Goodhew et al., 2013). Two major findings have emerged. First, the critical role that attention was thought to play in OSM has been cast into serious doubt. The original finding that set size interacts with masking magnitude (Di Lollo et al., 2000) was shown to be largely due to ceiling effects (Argyropoulos, Gellatly, Pilling, & Carter, 2013), suggesting that changes in attentional distribution have no impact on the magnitude of masking. When set size *was* found to interact with masking, this was shown to be due to crowding, rather than to attentional changes (Camp, Pilling, Argyropoulos, & Gellatly, 2015); and a recent study showed that masking remained unchanged when the area over which attention was distributed was manipulated (Goodhew & Edwards, 2016). Consistent with this, OSM was recently demonstrated in a fully attended and foveated target (Filmer, Mattingley, & Dux, 2015). Second, OSM has been shown to be reliably modulated by manipulating the degree to which the visual system is likely to

treat the target and mask as a single object. For example, Lleras & Moore (2003) showed that flashing the mask once at common onset with the target, and then again at a spatiotemporal interval conducive to apparent motion, produced masking; however, when the temporal interval was too long to produce apparent motion, masking was not obtained. This finding was extended in Pilling & Gellatly (2010), who showed that introducing display elements that interfered with the apparent motion between a target and a subsequently flashed mask reduced masking, suggesting that it is apparent motion that is the critical variable, rather than the temporal interval which had previously been confounded with apparent motion. Similarly, masking is stronger when the target and mask share similar features, for example color (Moore & Lleras, 2005), or gabor orientation (Goodhew et al., 2015). This second set of findings strongly points to a masking phenomenon that reflects an object *updating*, rather than a substitution process (Enns, Lleras, & Moore, 2010; Goodhew, 2017). Here, masking can be understood as exploiting the visual system's ability to maintain object continuity over brief changes or interruptions, such as those involved in saccades (Pollatsek, Rayner, & Henderson, 1990) and temporary occlusions (Burke, 1952). According to this view, when the initial representation (target plus mask) is changed to a mask only display, the visual system can either individuate these two representations, or can interpret the situation as involving a single object that has changed over time. In the former case, the target is successfully perceived, and in the latter only the mask is perceived. In a sense, this is reminiscent of Navon and Purcell's (1981) integration and interruption model (Chapter 1.2), where integration protected the target from being masked, and when conditions favoured target mask individuation, interruptive masking could occur. In object updating, however, individuation, or segmentation, protects the target from masking, and conditions that favour the two stimuli being treated as a single object allow

masking to occur. It is also interesting to reflect upon earlier work, where masking was stronger when the target and masks were similar, rather than different, in form (Experiment 4 in Uttal, 1970). In this experiment, a single shape was backward masked by two flanking shapes, and it is possible that similarity in form between the target and flankers was conducive to an interpretation that the target had moved to the left or right upon mask onset, thus reducing the visibility of the original target. The object updating account is a subtle but important distinction from the object substitution account. It is an exciting development that has placed OSM (which Goodhew, 2017, has suggested be renamed to object *segmentation* masking) within an established visuo-cognitive framework (Kahneman, Treisman, & Gibbs, 1992), and can be linked to other interesting visual phenomena such as repetition blindness (Goodhew, Greenwood, & Edwards, 2016).

One of the important features of OSM theory is that it is not necessarily incompatible with traditional masking accounts. That is, the existence of reentrant object updating mechanisms does not preclude the existence of other mechanisms, such as integration and spatially local inhibition. It does, however, challenge the idea that other theories, such as the sustained transient dual channel model (Breitmeyer & Ganz, 1976), can account for all masking phenomena (indeed, the authors of this dual channel model have never claimed such an idea), and it is certainly possible that some of the effects found in masking studies over the last several decades are at least partially due to mechanisms involved in OSM. In fact, the dual channel model has been updated to include reentrant processing (Öğmen, 1993)[4] and the interactions between the magno- and parvocellular systems have been hypothesized to occur within various hierarchies of

4 It should be noted, however, that Öğmen's (1993) model was developed, in part, to account for phenomena such as motion deblurring, and had nothing to do with object substitution.

reentrant processing across the visual system (Breitmeyer, 2007; Tapia & Breitmeyer, 2011; see also Kafaligonul, Breitmeyer, & Öğmen, 2015), where, for example, the magnocellular pathway is responsible for the initial rapid projection to higher cortical areas (Kveraga, Boshyan, & Bar, 2007). Nevertheless, it is clear that OSM is at least partially distinct from other masking phenomena (Breitmeyer, 2015). A particularly striking example of this comes from Chakravarthi & Cavanagh (2009), who found that while a common onset four-dot mask, a backward contour mask, and a backward noise mask all effectively and equally masked the flankers that were in position to crowd a target, only the latter two masks were effective at releasing the target from crowding. That is, while a four-dot mask reduced flanker identification to virtually chance level performance, it did not prevent the (masked) flanker from crowding (and thus reducing performance in identifying) the target. This is strong evidence that masking phenomena involving noise and metacontrast exist at a level of processing prior to that of crowding, while OSM occurs at a later stage.

1.6 Purpose of the current work

Until the discovery of OSM, the primary temporal parameter used to explore masking has been SOA (interstimulus interval (ISI), and stimulus termination asynchrony (STA) are close relatives that have also been used, e.g. see Macknik & Livingstone, 1998). Since the discovery of OSM, a different parameter—the duration of a common onset mask—has emerged, but has generally been reserved for studies involving masking with a sparse array of dots. This reflects the different mechanisms purported to underlie the respective forms of masking. When it comes to transient on sustained inhibition mechanisms, for example, SOA is precisely what determines the degree to which the magnocellular response of the mask is temporally superimposed

upon the parvocellular response of the target. In OSM, mask duration determines the likelihood that the target representation will be substituted or updated with that of the mask. Similarly, in OSM studies, the mask has typically comprised a set of dots, rather than more traditional mask types (e.g. an annulus), and this again reflects the mechanisms thought to be involved in OSM: the use of a dot mask is more likely to isolate OSM specific mechanisms, as its sparse configuration leaves little opportunity for spatially local contour interactions (Di Lollo et al., 2000; Enns, 2004). Thus, with few exceptions (Bischof & Di Lollo, 1995; Di Lollo et al., 1993; Di Lollo et al., 2000), there have not been systematic investigations into how masking changes as a function of mask duration with a spatially extensive mask, such as an annulus; nor has there been much work investigating different mask types within a common onset paradigm. Indeed, this methodological clustering is evident in the fact that the use of terms like metacontrast and OSM are often conflated with the spatial structure of the mask and the mechanisms thought to underlie the effect of each type of mask. Metacontrast technically means backward masking with a spatially adjacent mask (typically a mask whose contours are parallel with that of the target), but is sometimes used to describe the shape of the mask, independent of its temporal relationship with the target (e.g. Di Lollo et al., 1993), a point which is raised by Kahneman (1968). It is also a theoretically loaded term, evidenced by the term's origin (Chapter 1.1). This clustering is reinforced by the use of a multi stimulus array in parafoveal visual field in OSM studies, while backward masking studies generally use a single stimulus in central or parafoveal visual field. This restricted use of temporal and spatial parameters makes it challenging to compare the effects these parameters actually have upon masking. For example, it would be interesting to measure how the magnitude of masking compares between a backward and common onset mask when the spatial properties of the stimuli are kept constant.

Similarly, it would be interesting to compare the effects of different mask types within a common onset paradigm (for example, compare experiments 1 and 3 in Di Lollo et al., 2000). The experiments in this dissertation are a response to this observation. In the first set of experiments, a single stimulus type (faces) is used across a variety of temporal schedules, ranging from backward masking, to common onset masking, and other variations. In the second set of experiments, two different mask types (four-dot, and annulus) are used in a common onset paradigm. Importantly, in both sets of experiments, stimuli are presented foveally. In addition to eliminating eccentricity and attention as confounds, this also eliminates the influence of crowding, and means that eye movements do not have to be accounted for. It should be noted that after the first set of experiments were completed, Filmer et al., (2015) published their discovery of OSM in central visual field, and it was this discovery that prompted the second set of experiments.

In the first set of experiments (Chapter 2), successful masking was obtained with face stimuli in central visual field with both backward and common onset masking presentations. Masking was also obtained using other forms of mask schedules that were intermediate between these two forms. A simple computational model was developed to account for the data from these experiments. In the second set of experiments (Chapter 3), I successfully replicated Filmer et al. (2015), and obtained four-dot masking with a foveated target. Two more experiments were conducted to explore this further. First, masking was measured as a function of target mask separation. Here, four-dot masking remained relatively unaffected across a range of separations, suggesting an underlying spatially invariant mechanism. Second, a matching task was used to measure the magnitude of errors in orientation judgments using a four-dot mask. The key finding here was that masking appears to operate by rendering the target completely invisible,

rather than simply degrading it. In Chapter 4, I present a more sophisticated computational model that simulates several experiments across those described in Chapters 2 and 3.

Chapter Two

Masking With Faces In Central Visual Field

(adapted from Daar & Wilson, 2015)

SUMMARY

With a few exceptions, previous studies have explored masking using either a backward mask or a common onset trailing mask, but not both. In a series of experiments, we demonstrate the use of faces in central visual field as a viable method to study the relationship between these two types of mask schedule. We tested observers in a two alternative forced choice face identification task, where both target and mask comprised synthetic faces, and show that a simple model can successfully predict masking across a variety of masking schedules ranging from a backward mask to a common onset trailing mask and a number of intermediate variations. Our data are well accounted for by a window of sensitivity to mask interference that is centered at around 100 ms.

2.1 Introduction

The use of visual masking as a means to study the nature of visual perception has a long and rich history (Breitmeyer & Öğmen, 2006). By measuring the effect that varying spatiotemporal relationships between target and mask have upon visual processing of the target, valuable insights can be gained about the time course of visual perception (Bacon-Macé, Macé, Fabre-Thorpe, & Thorpe, 2005; Bar et al., 2006; Reeves, 1982), as well as spatial properties of vision (Ghose, Hermens, & Herzog, 2012; Habak,

Wilkinson, & Wilson, 2006). While masking continues to remain a popular tool in the study of vision (Breitmeyer, 2007), the general temporal character of the mask has been limited to two broad classes: those involving a briefly flashed (pulsed) mask (e.g. Burr, 1984) and those involving a common onset trailing mask (e.g. Di Lollo et al., 2000). In the former case, the primary temporal property of the mask that is studied is the stimulus onset asynchrony (SOA), where the onset of the mask is varied relative to the onset of the target (although there have been systematic investigations of the effect of changing the duration of target and mask pulses, e.g. Breitmeyer, 1978; Macknik & Livingstone, 1998). In the latter, the duration of the trailing mask is varied. The use of these classes of mask schedules has also led to the development of unique *spatial* relationships between target and mask structure. In most modern studies involving a pulsed mask, the contours of the mask and target are closely aligned, while studies involving a trailing mask typically use a sparse mask (Figure 2.1). Furthermore, pulsed contour masks are often studied in central visual field, while sparse trailing masks are usually studied in parafoveal visual field.

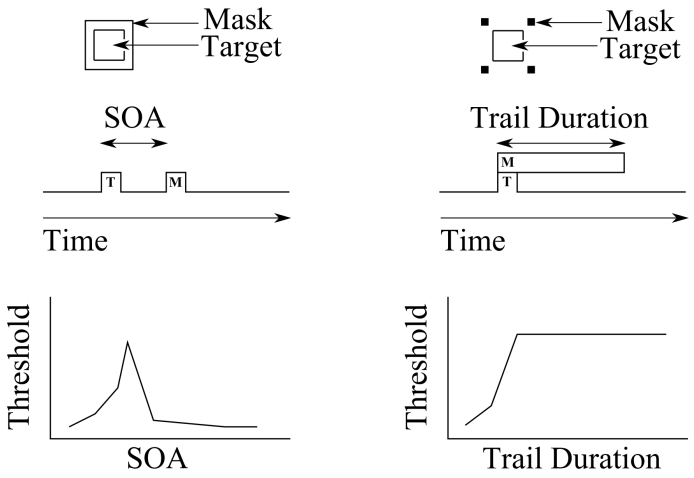

Figure 2.1. Two broad classes of masking schedules. **Left:** a pulsed backward contour mask. **Right:** a sparse four-dot common onset trailing mask. Typical masking functions for each are shown at bottom.

This packaging of stimulus properties is not accidental. The discovery that a sparse, four-dot mask can produce powerful masking supports the idea that the mechanisms of masking here involve interference with feedback from higher to lower areas of visual processing, as it is difficult to account for such masking with local feedforward effects such as lateral inhibition (although see Bridgeman, 2007). For example, the finding that robust masking can be obtained with non-foveal stimuli using large target-mask separations is difficult to explain with lateral inhibition (e.g. Growney, Weisstein, & Cox, 1977, although see Breitmeyer, Rudd, & Dunn, 1981). Accordingly, while a contour mask may derive its effectiveness through lateral inhibition and feedback (Enns, 2004), a sparse mask may be effective through feedback alone. Furthermore, the finding that it was, until recently (Filmer et al., 2015), challenging to

produce masking using a sparse mask with a single target in central visual field means that object substitution/updating studies often use multiple simultaneous targets arranged in parafoveal visual field. However, the fact that a sparse, trailing mask is generally ineffective in central visual field does not mean that the basic properties of object processing differ between central and parafoveal visual field. Object substitution masking (OSM) is thought to involve the interference of a masking pattern with feedback which, under normal (non masked) viewing, would serve to consolidate the target into conscious visual processing. The fact that OSM is not as effective in central visual field does not mean that feedback is not used to consolidate visual processing in central visual field; rather, this more likely means that these sparse masks are not powerful enough relative to the robust representation of information in central visual field (Enns & Di Lollo, 1997). In other words, feedback may still occur with central visual field representations, but these representations may be impervious to interference from a sparse mask. Importantly, this also implies that the effectiveness of a strong contour mask in central visual field may be at least partially due to object substitution mechanisms (Enns, 2004). Notably, a common onset contour mask in central visual field was found to produce powerful masking (Bischof & Di Lollo, 1995), showing that contour masking can occur without a delayed onset, and recent accounts of metacontrast masking include feedback mechanisms (Breitmeyer, 2007; Silverstein, 2015; Tapia & Breitmeyer, 2011).

One approach in exploring the extent to which mechanisms in backward masking and common onset masking overlap is to create a paradigm where the schedules of masking can be arbitrarily varied between the two extremes (backward pulse and common onset trail), while keeping constant both the spatial relationships between target and mask, and the location of presentation in visual field. In the current set of

experiments, we use centrally presented synthetic faces (Wilson, Loffler, & Wilkinson, 2002) for both target and mask, and explore the effects of varying mask schedule upon performance in a face identification task. The parametric generation of our faces allows us to titrate the difficulty of each testing condition to avoid ceiling and floor effects (Argyropoulos et al., 2013). Our particular faces are also interesting in that they contain elements of contour and camouflage masking (Figure 2.2).

This chapter is divided into three sets of experiments. The first set is designed to assess whether our stimuli can produce effective masking using both backward and common onset masking schedules, and to probe what effect, if any, briefly interrupting a trailing mask has upon performance. The second set follows up on this in order to determine if and when mask energy, across the duration of a single trial, has an additive effect upon performance. The last experiment follows up on the previous ones and explores the role of transients. Our results demonstrate the effectiveness of using faces in central visual field, both as a backward and common onset mask. Our results also suggest that additivity of mask energy across time depends upon the temporal window in question, and we have developed a model that suggests there is a distinct moment in time when the mask can interfere with target processing.

2.2 Experiment 2A

In this first experiment, we tested the effectiveness of our stimuli under two masking conditions: as a backward mask, and as a common onset trailing mask (Figure 2.3).

2.2.1 Stimuli

Synthetic faces (Wilson et al., 2002) were used for both target and mask patterns (Figure 2.2). The target face was either the mean of a set of 41 male faces, or a face whose distance from the mean was determined by a staircase procedure (two-down-one-up). On each trial, the identity of this latter face was randomly chosen from one of four orthogonal identities, and on each new run, four new identities were randomly chosen from the set of the 41 faces and then orthogonalized. The geometric difference between these faces and the mean face can be expressed as a percentage of geometric variation across the difference vector between the mean and non mean identity, relative to the mean head radius, and it was this geometric difference, or distance from the mean, that was varied according to the staircase procedure. The mean face was used as the mask, and was 50% larger than the target face. Our stimuli were presented on a VIEWPixx display, which has the advantage of a scanning backlight coupled with a fast pixel response time (pixel rise time and fall time are each 1 ms). The display was set to a refresh rate of 120 hz (8.3 ms per frame). We took advantage of these properties, and used an interleaved frame approach to present our stimuli. Presenting the target and mask in alternating frames is perceptually equivalent to the two stimuli being presented simultaneously at 60 hz at half contrast. At a viewing distance of 1.28 m, the screen subtended 23 by 13 degrees of visual angle, horizontally and vertically. The target faces subtended an average of 3.5 by 5.0 degrees, and the mask face subtended 5.25 by 7.5 degrees. The display was calibrated to linear light (gamma = 1), and the mean luminance of the screen, measured with a Konica-Minolta LS-110, was 59.8 cd/m^2.

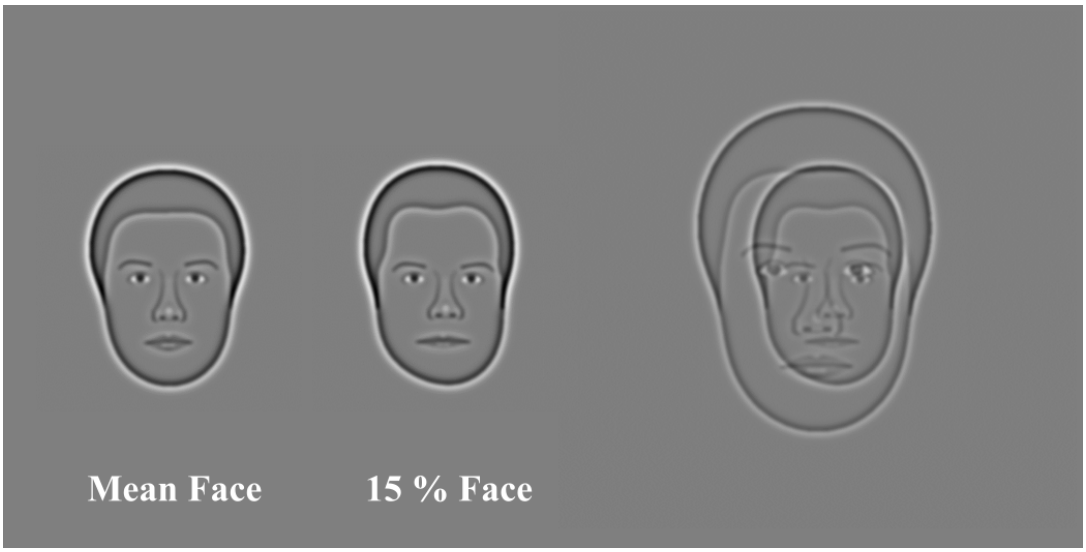

Figure 2.2. Target and mask stimuli. **Left:** mean face. **Middle:** face whose identity differs from the mean by 15%. **Right:** 15% face is masked by the mean face. In any given trial, the target face would be either the mean face, or one of a number of identities of various strengths. The mask, which was 50% larger than the target face, was always the mean face.

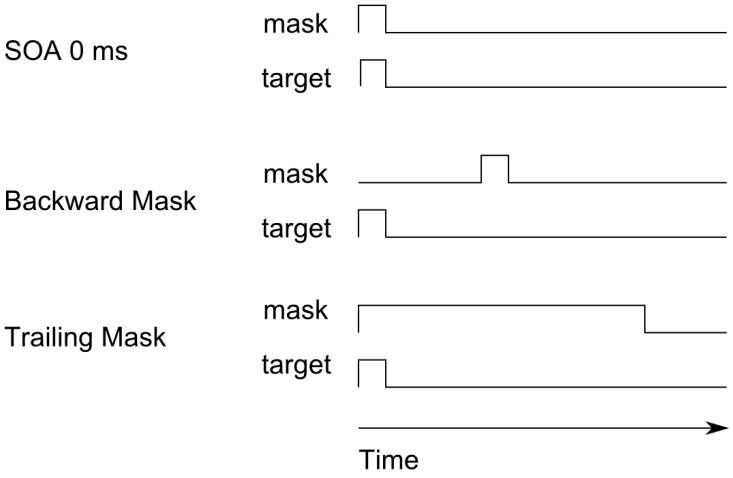

Figure 2.3. General schematic of backward and trailing mask conditions in Experiment 2A. The zero SOA condition is also shown.

2.2.2 Procedure

Figure 2.4 depicts the trial sequence. A keypress initiated each run. A white fixation cross at the center of a gray background appeared for one second, after which the fixation disappeared for 250 ms, followed by a briefly flashed target face (25 ms). In the backward masking conditions, a briefly flashed mask (25 ms) appeared at one of seven SOAs (0, 58.3, 75, 91.6, 125, 258.3, and 608.3 ms). In the trailing mask conditions, the mask onset simultaneously with the target onset (common onset), and trailed for one of seven durations that matched the SOA values (except for SOA = 0 ms). For each trial, the location of mask and target were independently spatially jittered by a random amount, up to a maximum of 0.61 degrees in the horizontal and vertical directions. A control condition was included where no mask appeared. After 1250 ms had passed from the beginning of the trial, a response screen displayed two faces (one of

which was always the mean), and the observer indicated with a keypress which of the two faces they thought matched the target face (which, on approximately half the trials, was the mean face). A soft tone indicated correct responses during training runs. No feedback was given for experimental runs. Including the mask absent control condition, there were a total of 14 unique conditions: seven SOA, seven trail, and the mask absent condition; as trail duration was calculated from the beginning of mask onset, the 25 ms Trail condition and the SOA 0 ms condition were one and the same condition. We tested four observers, three of whom completed two runs of each condition in random order; the fourth observer was unable to remain available for a complete set of runs, and was only able to complete one run of each condition, with the exception of Trail 608.3 ms, which was run twice. This experiment and all others reported in this dissertation were conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and were approved by the Research Ethics Board of York University. Informed consent was obtained from all observers prior to testing.

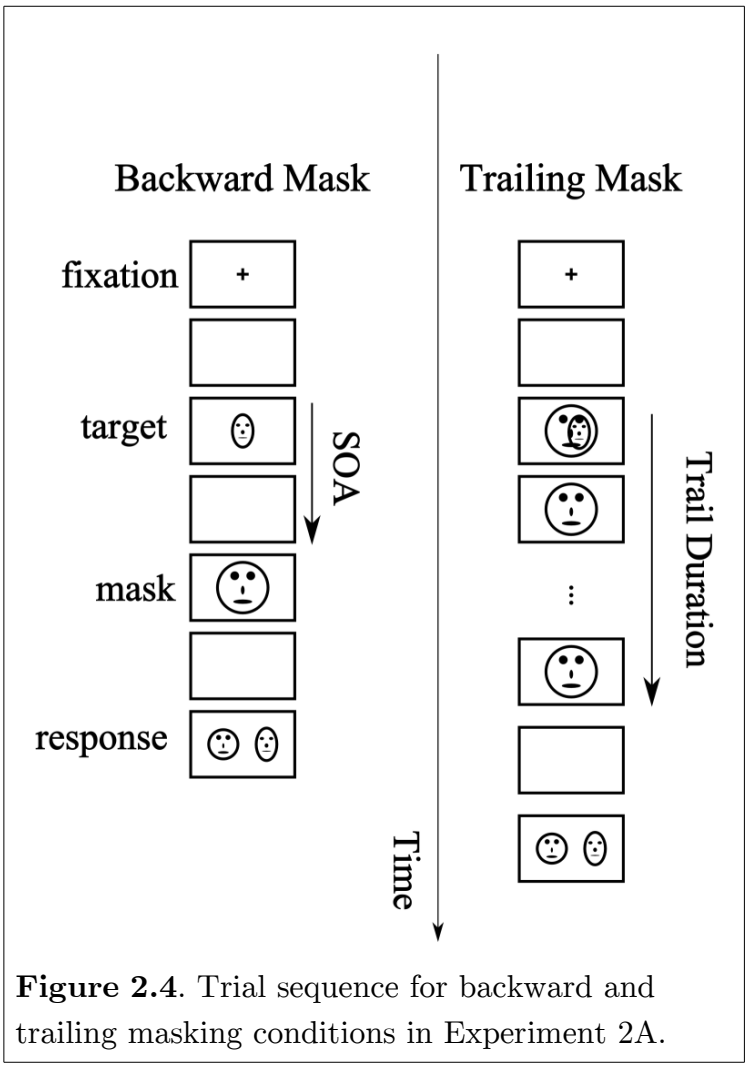

Figure 2.4. Trial sequence for backward and trailing masking conditions in Experiment 2A.

2.2.3 Results

Data from Experiment 2A are shown in Figure 2.5. For the backward masking condition (left), peak masking occurred at 58.3 ms. Data for the trailing mask condition are shown on the right. Masking did not appear to change as a function of trail duration, although we did not test at durations shorter than 58.3 ms. In both plots, the dashed line indicates the control condition (no mask), and in both conditions masking

elevated thresholds by a factor of about three. There was considerably more variation in the individual common onset masking functions compared to the backward masking functions, and this is reflected by the larger error bars in Figure 2.5.

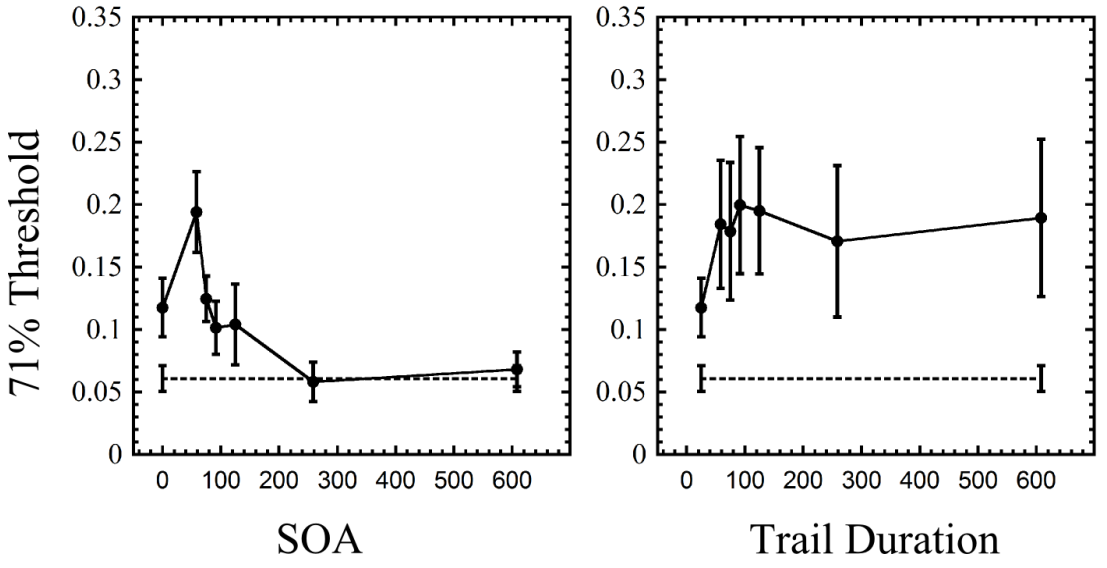

Figure 2.5. Data from four observers in Experiment 2A. Masking functions for both backward (**left**) and trailing (**right**) conditions are shown. Dashed line indicates unmasked threshold. Values indicate the face geometry difference between the two possible target faces (as a proportion) at which observers were able to perform with an accuracy of 71%, which is the percentage value that the staircase procedure converged upon. Error bars indicate Standard Error of the Mean across observers.

2.3 Experiment 2B

To investigate the relationship between the two forms of masking, we tested observers with an interrupted trailing mask. This was done by removing the mask for a

brief interval centered around the critical SOA (58.3 ms). If the masking found in the trailing condition was due entirely to the presence of the mask at 58.3 ms (see left panel in Figure 2.5), then removing it at this time point should eliminate masking. We tested seven observers in four main conditions (Figure 2.6). The first was an uninterrupted trailing condition identical to the 608.3 ms trailing condition in the previous experiment. Two interrupted trailing conditions were also run. The mask was removed for three frames (25 ms) in the first of these, and seven frames (58.3 ms) in the latter. As a control condition, the mask and target onset and offset together (SOA 0). This was chosen to control for the effect of any residual masking due to the overlap between target and mask in the interrupted trailing conditions.

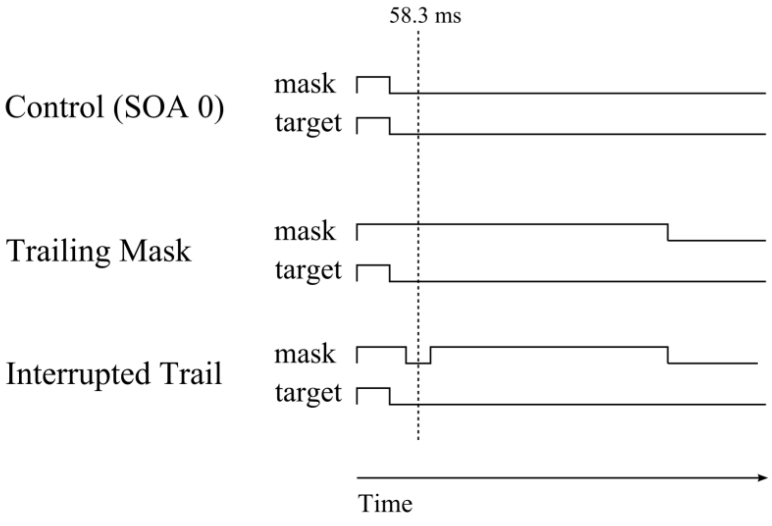

Figure 2.6. Schematic for conditions in Experiment 2B.

2.3.1 Results

Data from Experiment 2B are shown in Figure 2.7. The most notable finding here is that interrupting the trailing mask at this critical time period has no significant impact on masking. This finding held even when the mask was removed for seven frames (58.3 ms). A repeated measures ANOVA across all four conditions revealed a main effect of masking condition ($F(3,18) = 15.463$, $p < 0.0005$). Sidak corrected pairwise comparisons showed differences between the control condition (SOA 0 ms) and both the 25 ms wide gap and full trail ($p < 0.005$ for both comparisons), and a trend approaching significance between the control and 58.3 ms wide gap conditions ($p = 0.055$). Importantly, no difference was found between the three trailing conditions ($p>0.9$).

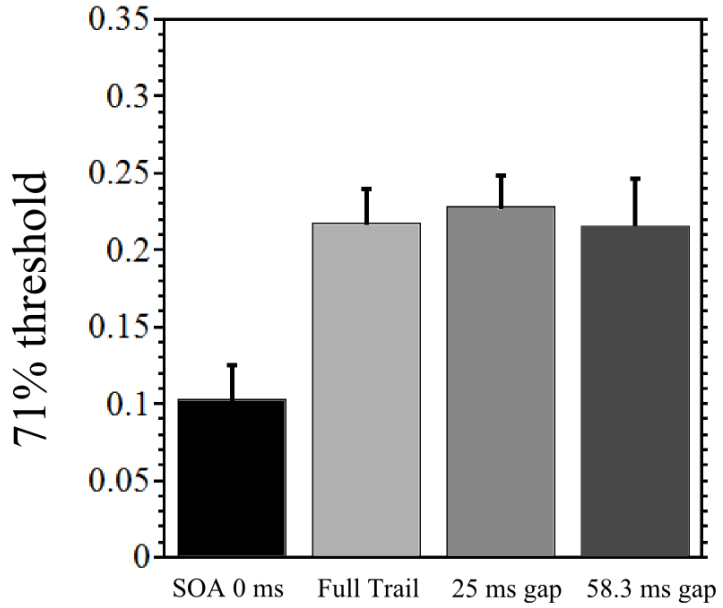

Figure 2.7. Data from Experiment 2B. Error bars indicate Standard Error of the Mean across observers.

2.3.2 Discussion

In our first set of experiments (Experiments 2A & 2B), we aimed to establish the viability of using centrally presented faces in both backward and trailing mask schedules. Our results confirm that these stimuli provide for effective masking, and reproduce patterns of masking found in previous studies of both backward (Burr, 1984) and trailing mask conditions (Di Lollo et al., 2000). In an attempt to determine whether the effectiveness of the trailing mask was due solely to its presence at the peak SOA (58.3 ms), we tested the effect of introducing a gap in the trail, centered at the peak SOA. We found that a gap as large as 58.3 ms (centered at 58.3 ms) did not reduce masking relative to the uninterrupted trail condition. Our next set of experiments were designed to investigate this finding in more detail.

2.4 Experiment 2C

The findings from Experiments 2A & 2B suggest that in the particular masking schedules tested, mask energy is not additive. If it were, then masking due to the uninterrupted trail condition should equal the sum of masking in the SOA 58.3 ms condition, and masking in the interrupted trail condition, and our data show otherwise. In order to explore this further, we created a set of conditions (Figure 2.8) specifically designed to test additivity across various combinations of mask schedules. As can be seen in this figure, the Gap condition represents the sum, in terms of mask energy across time, of the Early and Late conditions. If mask energy is additive, here, then the reduction in performance in the Early condition plus the reduction in performance in the Late condition should add up to the performance reduction in the Gap condition. The width of the gap in the Gap condition was 58.3 ms, and the width of the initial mask pulse in the Early condition was 50 ms. As a result, the gap was now centered

around 75 ms, rather than the 58.3 ms of the previous experiment. We did this to give us more room to work with to the "left" of this gap.

As we were measuring threshold elevations relative to the baseline condition (SOA 0), we decided to include a brief pulse at the start of the Late condition, matching the pulse in the baseline condition (all the other conditions naturally contained mask energy during those first 16.7 ms). This was done to ensure that in all conditions, any masking that was observed was above and beyond that produced by the mask being displayed at the same time as the target. As such, additivity is being assessed within a window that excludes the first 16.7 ms. We also included an uninterrupted trail condition to test whether a trailing mask would produce masking above that produced by an SOA of 0 ms. For the Full Trail, Gap, and Late conditions, the final mask offset occurred at 608.3 ms.

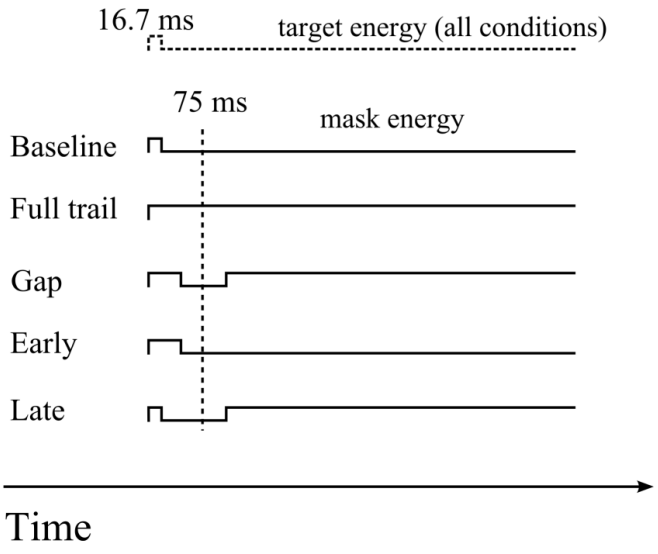

Figure 2.8. Conditions for Experiment 2C. The Gap condition represents the sum of the Early and Late components.

2.4.1 Procedure

The stimuli and procedure were similar to Experiments 2A & 2B, with a few exceptions. First, instead of adopting a staircase procedure, where the geometric distance between two target faces was gradually reduced as observers responded correctly, we chose, for each observer, a geometric distance that was at an appropriate difficulty level, and measured performance as proportion of correct trials. This choice of difficulty was accomplished during training where each observer completed several Full Trail masking trials at different difficultly levels until one was found where performance was between chance and ceiling. This enabled us to choose a custom difficulty level for each observer, and this level was constant for each observer across all five tested conditions. Nine observers completed these five conditions, and all five conditions were interleaved within each experimental run. Within each run, each condition was presented, in random order, 10 times. Each observer completed eight experimental runs. This meant that for each observer, performance at each condition was based on 80 trials. A second difference was that instead of presenting individual stimuli on alternating frames, we presented them simultaneously within each frame, allowing us to control the presentation of the stimuli with twice the temporal resolution. We reduced the initial target-plus-mask pulse to 16.7 ms (see Figure 2.8), and ensured that the contrast of all stimuli was kept constant throughout the sequence.

2.4.2 Results

To quantify performance reduction (relative to baseline), threshold elevations in any given condition were calculated as the proportion correct in the baseline condition divided by the proportion correct in the condition of interest. For example, if an observer was correct 60% of the time in a masking condition, and performed at 90% in

the baseline condition, the threshold elevation would be 1.5. Additivity was measured by comparing the partial threshold elevations. If mask energy in conditions A and B combine to form condition C, and mask energy is additive, then $(TE_A - 1) + (TE_B - 1) = (TE_C - 1)$, where TE denotes threshold elevation values. The data are shown in Figure 2.9.

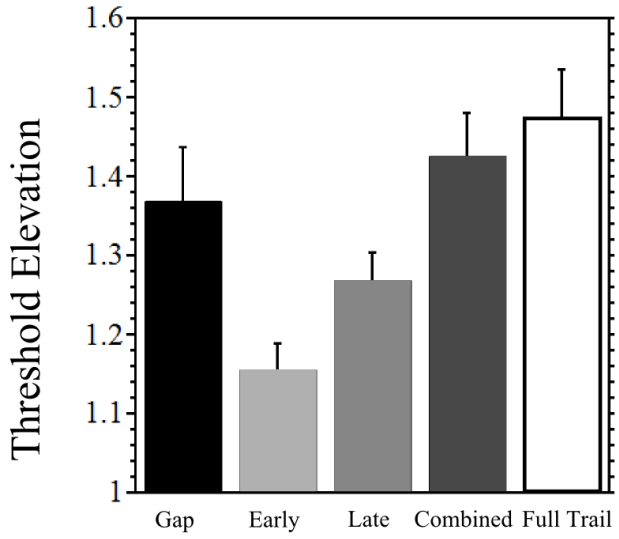

Figure 2.9. Data from Experiment 2C. The bar labeled "Combined" was calculated by summing the Early and Late data from each observer. Error bars indicate Standard Error of the Mean across observers.

The masking due to the Early component and the masking due to the Late component combine to form a value that is not statistically different from the Gap condition ($t(8) = 0.851$, $p = 0.419$); thus masking appears to be additive here. Our

results also demonstrate that masking in the Full Trail condition (white bar, black outline) produced masking beyond that produced in the S0A 0 ms condition: mean threshold elevation for the Full Trail condition was 1.47, which was significantly different from a value of 1.0 ($t(8) = 7.624$, $p < 0.0001$). As in Experiment 2B, the Gap condition did not significantly differ from the Full Trail condition ($t(8) = 1.195$, $p = 0.266$), although the Gap was centered around 75 ms and not 58.3 ms as in the previous experiment.

2.5 Experiment 2D

In Experiment 2D, we tested additivity in a different set of conditions, shown in Figure 2.10. To test for additivity, we asked whether masking in the Full Trail condition was equal to the combination of masking in the Gap condition plus masking in the Pulse condition (the Gap and Pulse conditions combine to form the Full Trail condition). Threshold elevations for the Gap and Full Trail conditions were used from the previous experiment, however a fresh baseline condition (SOA 0 ms) was included so threshold elevations could be calculated for the new Pulse condition. Data from the same nine observers as Experiment 2C are shown in Figure 2.11. Here, the masking due to the Full Trail is less than the sum of the masking in each of the components ($t(8) = 2.513$, $p < 0.05$). In other words, masking here is subadditive. It should be noted that across these additivity experiments, we have been assuming it is appropriate to simply add thresholds. A better approach may be to use a more linear measure such as d prime.

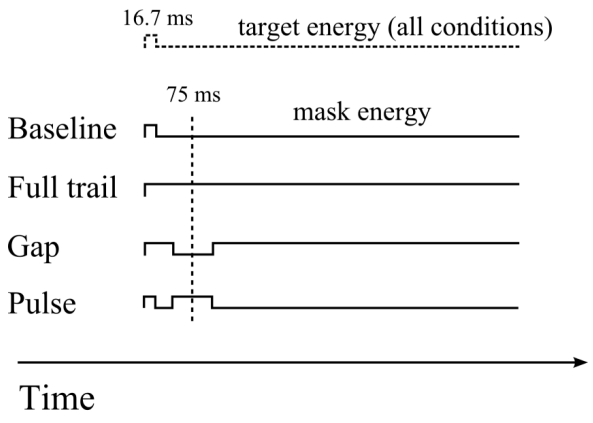

Figure 2.10. Conditions for Experiment 2D. The Full trail condition represents the sum of the Gap and Pulse conditions.

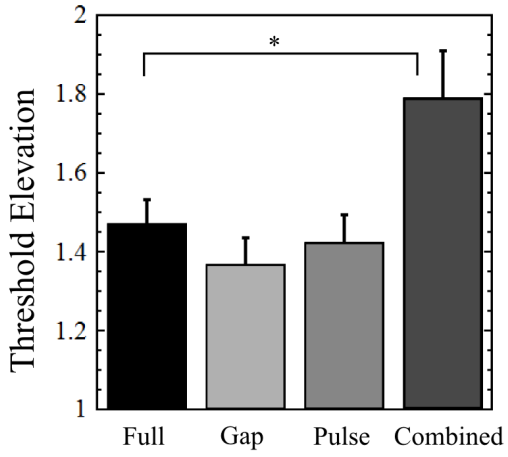

Figure 2.11. Data from Experiment 2D. Error bars indicate Standard Error of the Mean across observers.

2.6 Experiment 2E

In addition to the additivity experiments, we were curious to see whether we could obtain masking in excess of that due to a full uninterrupted trail, by repeatedly pulsing the mask throughout the duration of the trial. Our reasoning was that the transients associated with these mask pulses would perhaps provide greater interference with the consolidation of target representations. In Experiment 2E, two observers, one of whom was naive, were tested with eight different pulse frequencies, ranging from 12 hz (period = 83.33 ms) to 40 hz (period = 25 ms), in addition to a full uninterrupted trail (see Figure 2.12). In all conditions, the mask (whether it was pulsed or uninterrupted) lasted 608.3 ms. Each experimental run contained these nine conditions presented in random order, 10 times each. One observer completed eight runs (total of 80 trials per pulse frequency), and the other completed four runs (40 trials per pulse frequency) due to availability. Data are shown in Figure 2.12.

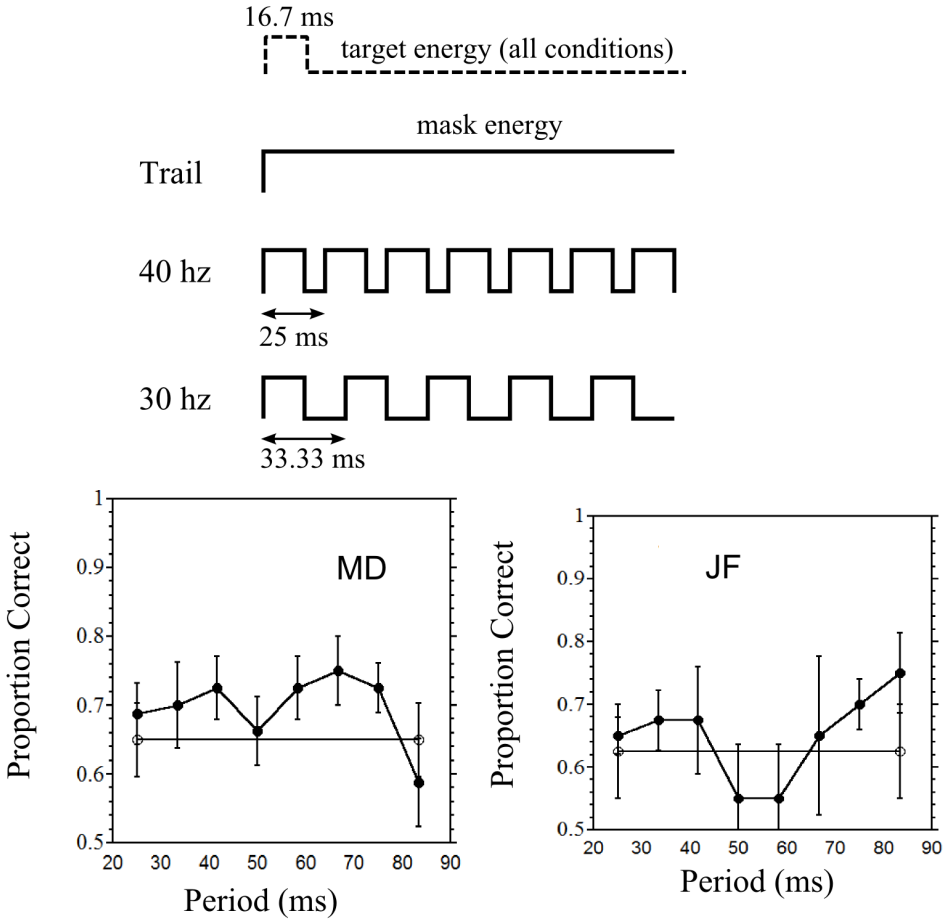

Figure 2.12. Three conditions from experiment 2E are shown (out of a total of 9). Starting from top: Full, uninterrupted trail; mask pulses that have a period of 25 ms (40 hz); pulses with a period of 33.33 ms (30 hz). Bottom: data for two observers as a function of period. The open circles indicate performance in the full trail condition. Note that lower values indicate worse performance. Error bars indicate Standard Error of the Mean across runs for each observer.

Neither observer showed an increased level of masking, relative to an uninterrupted trail (straight line, open circles), at any of the tested pulse frequencies. While it may be the case that masking is saturated with a common onset mask, a more likely explanation for this result is that the periodic stimulation resulted in a steady state response, rather than a series of individual transient responses (see Norcia et al., 2015 for a review). Even the lowest frequency used in Experiment 2E (12 hz) is fast enough to produce a steady state response, so it is entirely possible that the cascade of events associated with a single transient was not present with the periodic stimulation. It would be interesting to test more observers with such periodic stimulation. It could be the case, for example, that there are interesting idiosyncratic differences between observers. Alternatively, certain frequencies may produce reliable effects across a large observer pool.

2.7 Computational Model

The finding that mask energy, across time, is additive in some scenarios but not others is intriguing, and to explore this further, we developed a computational model to account for our data. Based on recordings of single cell responses to stationary gratings (Camp, Cheong, Tailby, & Solomon, 2011), activity in the magnocellular channel is simulated using the following equation, which models the transient and sustained component of neural activity to a stimulus:

$$M(t) = S + (T - S) \cdot e^{\frac{-(t - t_0)}{\tau}} \qquad \textbf{Equation 2.1}$$

where M(t) is the magnocellular response of the mask, S is the amplitude of the sustained level of this response (1.8% of the initial amplitude), T is the amplitude of the

transient response, and τ is the decay constant (39 ms). This equation, and the values for S/T and τ are taken directly from Camp et al. (2011). It should be noted that in this context, the terms transient and sustained do not refer to magnocellular and parvocellular cells, respectively, but rather to the magnocellular response to transient and sustained components of the *stimuli*.

In our model, masking occurs when the neural activity in magnocellular channels generated by the mask inhibits, or otherwise interferes with, a distinct neural process. This window of vulnerability, over time, is modeled as a Gaussian, G(t). The response to the target is modeled as follows:

$$R_T = C_{Targ} \cdot A \cdot \int_0^{608} \frac{G(t)}{1 + G(t) \cdot M(t)} dt$$ **Equation 2.2**

where A is a constant, and C_{Targ} is the contrast of the Target stimulus, nominally set to a value of 1.0 here.

Masking was modeled by taking the reciprocal of R_T, and feeding this value into the following contrast gain function, which Tapia & Breitmeyer (2011) used to model the contrast response of magnocellular cells in primate retina (Kaplan & Shapley, 1986).

$$Masking = \frac{R_{max} \cdot C^{\alpha}}{C_{0.5}{}^{\alpha} + C^{\alpha}}$$ **Equation 2.3**

where R_{max} is the maximum response of the function (set to 3.5, see below), C is the reciprocal of R_T, $C_{0.5}$ is the semi saturation constant, and alpha controls the slope of the

function. $C_{0.5}$ and alpha were set to values of 0.13, and 1, respectively, which are the values Tapia & Breitmeyer (2011) used in their model.

Our model contained four free parameters: The amplitude, mean, and standard deviation of G(t), and the constant A (Equation 2.2). Data from all unique conditions from Experiments 2C & 2D were used to generate the best fitting parameters. In addition, data from three conditions in Experiment 2A were included: the completely unmasked condition, SOA 58.3 ms, and SOA 258.3 ms. As different baselines were used across these experiments, all data were scaled such that threshold elevations were measured relative to the completely unmasked condition. As a result, the common onset common offset condition now has a threshold elevation of 1.938, and the largest threshold elevation value was 3.2, for the SOA 58.3 ms condition. Given the data across all our experiments, we chose an upper limit of masking at a threshold elevation of 3.5, which accounts for our choice of the R_{max} parameter in Equation 2.3. The fitting procedure was a simple minimization of the residual between the predicted and actual threshold elevations, across all nine masking conditions. The results are shown below (Figure 2.13).

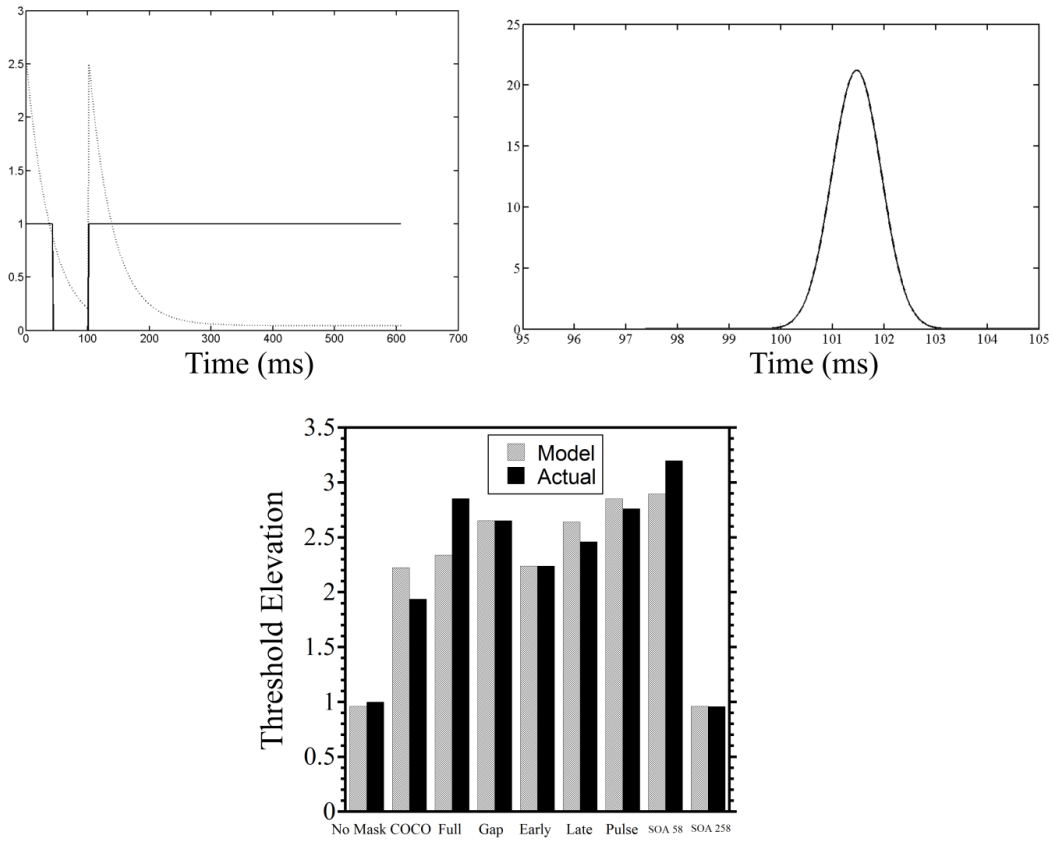

Figure 2.13. **Top Left**: Energy of mask stimulus (solid line) overlaid with modeled magnocellular response to this stimulus (dotted line) for the Gap condition. **Top Right**: Model output for G(t). **Bottom**: model predictions vs. actual data across the nine masking conditions. COCO = common onset common offset. Parameter outputs are as follows: Mean of G(t) = 101.5 ms; Standard deviation of G(t) = 0.47 ms; Amplitude of G(t) = 41.12; A = 26.40.

Our primary motivation in running Experiments 2C & 2D was to more closely investigate the finding from Experiment 2B that interrupting a trailing mask, during the peak SOA, did not appreciably diminish masking. Our model can account for this

finding, as when stimulus energy offsets, for example at the beginning of the brief gap in an interrupted trail condition, there is a lingering neural response that has potential to interfere with G(t), even though the mask stimulus is physically absent (see Figure 2.13, top left, which depicts an interrupted trailing mask). The reason that masking in the Gap and Pulse conditions didn't combine additively to resemble masking found in the Full Trail condition (Figure 2.11) is that our model has an upper limit to masking.

In our model, we did not make explicit what G(t) represents, other than the idea that it is crucial for the conscious perception of the target. One possible interpretation is that it represents reentrant parvocellular activity of the target, which is susceptible to inhibition by mask modulated magnocellular activity (Breitmeyer, 2007). While the temporal location of G(t) is consistent with a feedback interpretation (Fahrenfort et al., 2007), our data alone cannot distinguish whether G(t) represents feedback, feedforward, or some combination of the two.

Established models of metacontrast masking (e.g. see Breitmeyer 2007) purport two distinct mechanisms of masking: intrachannel inhibition, where the parvocellular response of the mask inhibits the parvocellular response of the target, and interchannel inhibition, where the magnocellular response of the mask inhibits the parvocellular response of the target. We initially modeled parvocellular activity of both the target and the mask, and included intrachannel inhibition as a contributor to masking; however, the output of our model rendered the intrachannel component of masking virtually irrelevant, as it contributed to about a thousandth of the total masking. For this reason, we decided to only include the interference of the magnocellular component of the mask with G(t). The contribution of two distinct masking components, each subserved by channels with different contrast gain functions, elegantly explains the shift from type B to type A backward masking functions as the ratio of mask to target energy increases

(Breitmeyer, 1978). When the strength of the mask increases relative to that of the target, the intrachannel component of masking becomes proportionally larger, relative to the interchannel component, and the particular way in which backward masking functions transition from type B to type A can be neatly accounted for by the particular contrast gain functions of magnocellular and parvocellular cells. While our model doesn't include these two components, it nevertheless can model a similar transition (Figure 2.14). In our case, however, the shift can be accounted for by the explicit inclusion of an upper limit on masking, rather than a shift from interchannel to intrachannel inhibition.

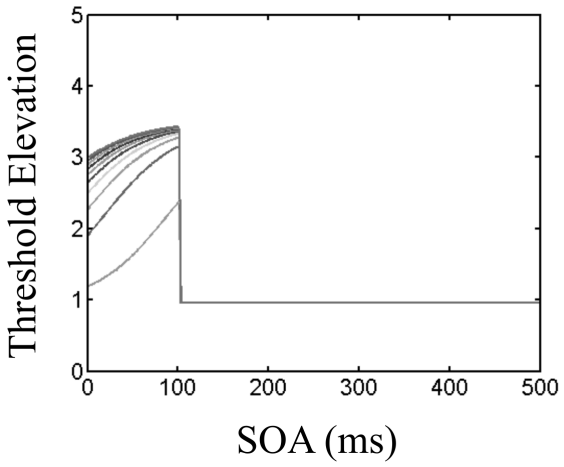

Figure 2.14. Transition from type B to type A backward masking. Shown here are backward masking curves, across 10 different mask contrasts, ranging from 0.1 to 4.6. As mask energy increases, the peak of the function becomes less pronounced, as masking at lower SOAs increases relative to that at higher SOAs.

Our finding in Experiment 2E suggests that it is unlikely that mask transients introduced after the initial common onset pulse can increase the amount of masking compared to that obtained with a full uninterrupted trail. As discussed in Chapter 2.6, however, this may reflect a steady state, rather than a transient response to the periodic stimulation. It's also worth noting that the higher frequency masking pulses may have been too rapid to be individually encoded by the magnocellular pathway (Poggel, Treutwein, Calmanti, & Strasburger, 2006). Coupled with the fact that only two observers ran this experiment, the results of Experiment 2E do not definitively clarify the potential role of transients subsequent to the initial mask onset. To test this further, we ran a final experiment that explored how introducing luminance transients on top of a trailing mask would affect masking.

2.8 Experiment 2F

One of the key findings from the previous experiments is that introducing a gap in a trailing mask does not reduce masking, even though this gap is centered around the SOA at which peak masking occurs in a backward masking condition. This suggests that the presence of the mask at this time period is sufficient but not necessary for strong masking. However, as was suggested by an anonymous referee, it is possible that the removal of the mask during this period does indeed reduce masking, but that the onset transient involved with the reintroduction of the mask at the end of this gap produces its own masking, and together, these effects cancel out. To further explore the role of transients, we ran five observers in a new set of conditions, which introduced transients during the peak SOA, during a 33.3 ms window (Figure 5.15). In addition to the **Full Trail** and **Gap** conditions, we added a third condition (**1 Pulse**), where the mask doubled in contrast during the 33.3 ms window. If this condition produced

masking above and beyond the **Full Trail** condition, then this would be compatible with the idea that transients can have an effect within a common onset masking context. Our model would also predict an increase in masking under this condition (up to a point, due to the saturation limit in our model), simply by virtue of there being extra masking energy to interfere during the window of sensitivity. On the other hand, if no additional masking occurs, this would suggest that (later) transients do not play a role in a common onset context (e.g. due to the common onset mask transient saturating the magnocellular pathway at start of trial). A fourth condition was also included (**2 Pulse**) where two double contrast mask transients were pulsed for the first and last 8.3 ms of the 33.3 ms window, with zero mask energy between them. If mask energy alone is what matters, then this condition should produce masking equivalent to the **Full Trail** condition, as the product of luminance and time is equivalent for these conditions.

The stimuli and methods were similar to Experiments 2C-E, with two main exceptions. In the previous three experiments, while the difficulty level (expressed as the percent difference in geometry between the two possible target faces) remained the same within each observer's set of runs, the particular identity of the non-mean face changed from trial to trial, and block to block. To reduce the variability in performance associated with different identities, a single identity was used throughout Experiment 2F, across all observers. The second change was that a preliminary set of backward masking trials were run on each observer to assess their peak SOA. Four SOAs were tested: 0 ms, 58.3 ms, 75 ms, and 250 ms. The results from these trials were used in the main experiment, so that each observer had the gap or transient centered on their own individual peak SOA (either 58.3 ms or 75 ms). A sixth observer showed a type A

masking function, with the strongest masking occurring at an SOA of 0 ms, and was excluded from the main experiment. For each condition, including each of the SOA values, observers ran a minimum of 90 trials.

2.8.1 Results

Data are shown in Figure 2.16. A repeated measures ANOVA showed a significant main effect of (common onset) masking condition ($F(3,12) = 6.671$, $p < 0.01$). Follow up planned comparisons showed a significant difference between The Full Trail and the 1 Pulse conditions ($t(4) = 2.951$, $p < 0.05$), no difference between the 1 Pulse and 2 Pulse conditions ($t(4) = 1.920$, $p = 0.127$), and a trend towards significance in the difference between the Full Trail and 2 Pulse conditions ($t(4) = 2.153$, $p = 0.098$).

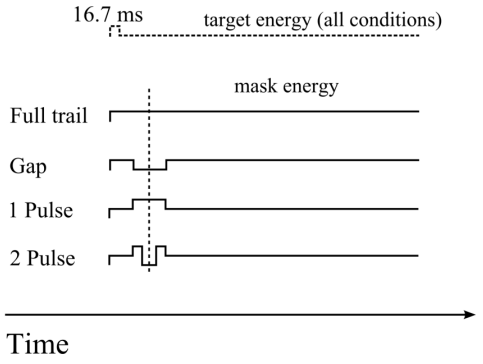

Figure 2.15. Conditions for Experiment 2F. In the Pulse conditions, the mask briefly doubles in contrast relative to the rest of the mask trail.

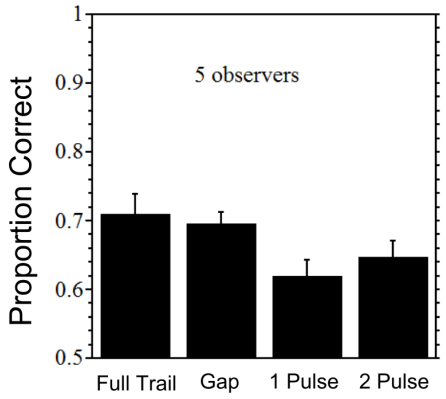

Figure 2.16. Data from Experiment 2F. The bar graph shows the means of the five observers who ran in the common onset conditions.

2.8.2 Discussion

In the current experiment, we have shown that a luminance transient on top of the trailing mask is able to further drive masking. While this result does not answer the deeper question of whether it is the transient itself or the stimulus mask energy during that period that is responsible for masking, it does suggest that masking with a full trail does not, in fact, represent a saturation of masking (at least with the mask and target contrasts used in our experiments).

The finding that the 1 Pulse and 2 Pulse conditions did not differ significantly from each other suggests that the mask energy of the physical stimulus alone does not

account for masking strength. Both these conditions had transients that occurred around the peak SOA, yet the difference in mask energy was identical to the difference between the Full Trail and 1 Pulse conditions. If stimulus mask energy during the critical SOA is the only thing that drives masking, then one would expect the 1 Pulse condition to exhibit significantly more masking than the 2 Pulse condition. Our model, however, predicts very little difference between these conditions, as it is not the mask energy of the *stimulus* that drives masking, but the energy of the *representation*, which incorporates a decay. Another possibility that should be noted is that the 2 Pulse condition contained on and off transients that were very close together in time. It may be the case that reciprocal inhibition between the on and off transients reduced the effectiveness of the pulses (Phillips & Singer, 1974; Singer & Phillips, 1974). This could account for the finding that the 2 Pulse condition showed a trend towards weaker masking compared to the 1 Pulse condition. Such an explanation might also account for the fact that masking was not shown to increase in the high frequency pulse conditions in Experiment 2E. Even if the steady state response can be modeled as a superposition of individual transient responses (Capilla, Pazo-Alvarez, Darriba, Campo, & Gross, 2011), the transient responses themselves may have effectively cancelled each other out through reciprocal inhibition.

While the results from Experiment 2F do not fully clarify the role of transients, they do show that our Full Trail condition does not saturate potential masking. The pattern of results we found is also qualitatively consistent with our model.

2.9 General Discussion

The experiments described here demonstrate the use of faces in central visual field as a viable tool for studying masking. Using a backward mask, we obtained peak

masking at a non zero SOA (type B masking). Using a trailing mask, we obtained masking greater than that with a common onset common offset mask (SOA 0 ms), demonstrating that the trailing portion of the mask was responsible for this effect. Our model, which was based on data from common onset masking with various portions of the trailing mask removed, as well as from backward making data, suggests that there is a specific period during the time course of visual processing, when the presence of a mask can exert itself. Our model also provides a framework within which backward and common onset masking operate through the same mechanisms. In particular, the notion of persistence plays an important role here, as it allows a decaying representation of a mask to have an impact on masking even when the physical mask has disappeared, and can therefore account for strong masking even in the presence of a gap (e.g. Experiment 2B). In fact, this very same idea of a persisting mask representation was incorporated by Jannati, Spalek, & Di Lollo (2013). While our Experiment 2B tested a common onset mask with a gap in the trailing mask, they used a converse configuration, where a common onset mask was followed by a single pulse. Their finding of strong masking with this configuration was explained by the decaying representation of the common onset mask being boosted by the following pulse.

While our model provided good fits to our data, there are some important limitations. For one, it does not take into account offset transients, which have been shown to be associated with threshold elevations, at least in the context of masking by light (Crawford, 1947). The account given by Macknik & Livingstone (1998), who used bars as targets and mask, suggests that the disinhibitory rebound associated with a sharp luminance decrement of the mask (offset transient) is responsible for inhibiting the target. As our model does not take offset transients into account, it could be underestimating the masking effect of the offset transients introduced when a trail is

interrupted, and it would be interesting to follow up this work to more closely explore the potential role of transients in our paradigm (e.g. Sackur, 2011; Tapia, Breitmeyer, & Jacob, 2011).

Our model also predicts that backward masking will peak at an SOA of around 100 ms, whereas our data suggest a peak quite a bit earlier (our baseline SOA runs in Experiment 2F showed a peak SOA of 58.3 ms for a few participants). This discrepancy could be due to any number of the following reasons. First, our model was based on data from methodologically different experiments. While the common onset data was derived from the same observers, using the same method of calculating threshold elevations, the backward masking data in Experiment 2A was derived from a different set of observers, using a different method of calculating threshold elevations. Indeed, when combining the two data sets into a unified set of elevation thresholds, the peak SOA condition from Experiment 2A showed slightly greater masking than the Full Trail condition from Experiment 2C, while within Experiment 2A, the peak SOA thresholds and Full Trail thresholds were virtually identical. Second, our model is a rather simple one, in that it posits a single integration across the entire trail duration, while in reality, there may be a number of distinct mechanisms each of which has its own integration window (see Breitmeyer & Öğmen (2006), p. 50). Third, our model only takes into account the magnocellular response of the mask, and does not model the influence of mask modulated parvocellular activity upon target visibility. It should also be noted that our model predicts unusually sharp backward masking functions (see Figure 2.14), which is due to the narrow width of G(t).

Another important issue is that it is challenging to determine whether the increased masking found with a common onset trailing mask, relative to a common onset common offset mask, is due to the trailing portion's ability to interfere with target

processing for an extended period of time, or whether it is simply due to the extended trail being temporally integrated into a (perceptually) higher contrast mask. In other words, increasing mask duration may be equivalent to simply using a higher contrast mask that offsets with the target offset, within a limited integration window (Bloch's Law). Increasing mask contrast relative to that of the target, while keeping mask and target durations fixed, has been shown to increase masking at an SOA of 0 ms (Stewart & Purcell, 1974; Weisstein, 1972). Similarly, increasing mask duration, while keeping mask and target contrast fixed, also results in increased masking at an SOA of 0 ms (Breitmeyer, 1978). Thus, it is difficult to say whether the increased masking we found with a common onset trailing mask is due to stronger sustained-on-sustained inhibition, or whether it is due to the trailing mask having more *time* to interfere with target processing (our model only takes into account the latter possibility). Di Lollo et al. (2000) found common onset masking with brightness matched masks, although that study involved sparse masks presented in the parafovea. There is, however, some interesting data that suggest that the duration of a centrally presented contour mask is capable of driving masking independently of its (luminance x time) energy. Di Lollo, von Mühlenen, Enns, & Bridgeman (2004) showed that while masking remained fairly constant across brightness matched targets of varying durations (increasing target duration or increasing target luminance reduced masking substantially), the same was not true when it came to mask energy. That is, increasing the duration of a brightness matched mask produced a very similar masking function to that when the mask duration was increased with a fixed luminance, suggesting that mask *duration* is capable of driving masking (see Bischof & Di Lollo, 1995, who also used brightness matched masks). While this wasn't common onset masking, the mask manipulations were done with a target duration of 10 ms, and an ISI of 0 ms, which is very close to a common

onset paradigm. It should be noted, however, that in these experiments, the mask was a contour mask, whereas in our experiments, the mask was a full face. Because there is more contour overlap with the target in a full face mask, there may be a larger component of intrachannel sustained masking, or masking due to integration (luminance channels of the mask and target being shared due to spatial overlap), relative to a simple contour mask. As such, mask energy (luminance x time) may have a larger role to play with our stimuli.

Our framework in exploring these various masking schedules shares similarities with the object substitution/updating framework, in that our model presupposes a scenario where the target representation is vulnerable to interference over a given period of time, and that this scenario can account for masking under a variety of temporal schedules. It is not clear, however, that sparse masks in parafoveal visual field operate through mechanisms that completely overlap those involving metacontrast masks in central visual field. For example, suppressing the magnocellular pathway by presenting stimuli on a red background attenuates metacontrast masking (Breitmeyer & Williams, 1990; Edwards, Hogben, Clark, & Pratt, 1996), while saturating the magnocellular pathway with a pulsed luminance pedestal has been shown to increase object substitution masking (Goodhew, Boal, & Edwards, 2014). This dissociation points to distinct masking mechanisms in metacontrast and object substitution. It should be pointed out, however, that in the latter study, the reduction in masking was measured as the performance change between baseline (a common onset common offset mask) and a common onset trailing mask that lasted 160 ms. If, as the authors argue, OSM reflects a failure of the visual system to temporally segment objects, then reducing the system's ability to temporally segment these objects (by pulsing a luminance pedestal) should result in a decreased performance in the trailing mask condition. In fact, in Experiment

2 of their study, there was no significant difference between the pulsed and steady pedestal conditions for either the baseline or the trailing condition. Rather, only an interaction was found to be significant. Moreover, inspection of their data suggests that the bulk of this effect was due to the baseline condition being easier in the pulsed pedestal condition, while the trailing condition showed hardly any difference between the pulsed and steady pedestal conditions (Figure 4 of their study). Thus, it is not clear that their manipulation actually increased masking via their proposed mechanism.

A second way in which metacontrast masking can be compared to OSM (or, rather, to common onset four-dot masking) is by how each is modulated by attention. There is evidence that metacontrast masking can be modulated by attention (Boyer & Ro, 2007; Ramachandran & Cobb, 1995). On the other hand, the role of attention in four-dot masking is less clear. A recent series of studies have investigated whether set size interacts with the magnitude of OSM (Argyropoulos et al., 2013; Filmer et al., 2014; Camp et al., 2015). If, as originally claimed by Di Lollo et al. (2000), larger set sizes result in increased masking, then this would suggest that attention modulates masking, as larger set sizes presumably reduce the amount of attention that can be allocated to the target, and/or the time it takes for the target to be attended to. The most recent of these studies (Camp et al.) provides convincing data that this interaction exists, however their Experiment 4 elegantly dissociates set size from *crowding* (which was previously confounded), and makes a strong case that it is in fact crowding that interacts with masking, rather than set size per se. Crowding, however, is not essential for OSM. A recent study demonstrated common onset masking with a four-dot mask in central visual field by degrading the target representation with a forward noise mask (Filmer et al., 2015), a finding that we have recently replicated in our own lab (Daar & Wilson, 2016; see Chapter 3).

Another set of experiments investigated whether spatially precuing the target location would reduce OSM (Pilling, Gellatly, Argyropoulos, & Skarratt, 2014). The bulk of their data revealed that spatial precuing did not decrease OSM. That is, while the manipulation did improve performance, it did so equally across all mask durations. Taken together, these studies yield important insights into OSM. Perhaps most important is the finding that some manipulations can degrade overall performance (but without increasing masking), while other manipulations selectively degrade performance at higher mask durations. It may be the case that there is a threshold for the quality of a target representation (relative to that of the mask), at which OSM occurs, whereas if the target is able to survive the iterative process of reentrant processing, its perceptibility is modulated depending upon the conditions of any given trial. Thus, if the target representation is degraded, but the threshold is not attained, one would expect an overall performance reduction without an increase in masking, whereas if the threshold is met, the target simply never enters into awareness, degraded or otherwise. Another possibility is that there are certain types of stimulus manipulation that, by their very nature, do not selectively degrade performance at higher mask durations—in other words, these manipulations do not interact with mask duration. For example, pilot work from our lab suggests that, when trying to obtain four-dot masking in central visual field, it is possible to have a baseline performance (SOA 0 ms) below ceiling by reducing target contrast, yet to observe no reduction in performance with longer mask durations. However, when a forward noise mask is introduced, as in Filmer et al. (2015), reliable masking occurs. If these data bear out, then this would suggest that the difference in the quality of a noise masked target, compared to a low contrast target, is relevant insofar as OSM is concerned. Further work in both metacontrast and common onset masking will be needed to clarify these possibilities. The paradigms employed in

Kahan & Enns (2010), and Harrison, Rajsic, & Wilson (2016) may prove especially useful here.

Mask preview (Tata & Giaschi, 2004; Lim & Chua, 2008) offers other intriguing insights into the role of attention in OSM. Tata & Giaschi found that when the masks were previewed for a brief period, OSM was virtually abolished (this could not be attributed to inadvertently cuing the target location, as all masks in the array were previewed). This result is neatly accounted for as follows: previewing the mask means that the visual system has taken care of attending to the mask in advance of the common onset phase, thus allowing the target to be attended without competition from the mask. In a fascinating follow up to this work, Lim & Chua discovered that this mask preview effect could be eliminated if the previewed mask and the following mask were displayed in such a way that the visual system was likely to treat them as distinct objects. Thus, it appears that attention plays a significant role in OSM, and that it likely operates largely on an object level representation. Our model is thus limited in two ways here. First, it does not say anything about object continuity of the stimuli, and can therefore not predict what happens under manipulations which, for example, either preserve or abolish the continuity of a mask across the trial (e.g. Lleras & Moore, 2003). Second, the model is not able to predict what happens under the conditions of mask preview. To incorporate mask preview, a term could be introduced that scales the response of the mask, M(t), such that in the extreme, a perfectly effective mask preview will abolish masking by reducing M(t) to 0. In order to model this, data would be needed to see what sort of scaling function is appropriate, or whether indeed the phenomenon is best described dichotomously.

In our experiments, we used faces in a two alternative forced choice task, whereas many other studies use more basic stimuli (e.g. a Landolt C) where the task is

to determine where the gap is, rather than to choose from alternatives that are explicitly presented at the end of each trial (although one could conceivably present the alternative Landolt stimuli at the end of such a trial). The use of different target and mask stimuli has been shown to have a substantial effect on the shape of backward masking functions (Francis & Cho, 2008), so our particular results may not be fully generalizable to other combinations of stimuli, as discussed above. The use of different task parameters and different criterion contents has also been shown to have an impact on masking functions (Breitmeyer & Öğmen, 2006), and it is possible that the complexity of our stimuli could have resulted in different strategies between and within observers. For example, using figural properties of the features as a basis of discrimination may produce different functions than using the contours of the head outline (when questioned, all of our observers reported that the head outline of the target was the most reliable cue for discrimination). While our face stimuli did produce powerful masking in central visual field, and have been used by our lab previously to study backward masking (Loffler, Gordon, Wilkinson, Goren, & Wilson, 2005), it would be useful to compare these results to more basic, and commonly used stimuli.

Our current experiments do not prove that the common onset masking we obtained was due to object substitution mechanisms, and we do not claim that our model captures the full complexity and array of mechanisms involved in masking. It does, however, provide new data that can be directly assessed in the light of various models and masking mechanisms, while introducing some novel forms of mask presentation. Future experiments that build upon these ideas may be able to further unravel the relationships between backward and common onset masking.

Chapter Three

A Closer Look At Four-Dot Masking

(adapted from Daar & Wilson, 2016)

SUMMARY

Four-dot masking with a common onset mask was recently demonstrated in a fully attended and foveated target (Filmer et al., 2015). Here, we replicate and extend this finding, by directly comparing a four-dot mask with an annulus mask while probing masking as a function of target-mask separation. Our results suggest that while an annulus mask operates via spatially local contour interactions, a four-dot mask operates through spatially global mechanisms. We also measure how the visual system's representation of an oriented bar is impacted by a four-dot mask, and find that masking here does not degrade the precision of perceived targets, but instead appears to be driven exclusively by rendering the target completely invisible.

3.1 Introduction

Over the last two decades, common onset masking with a four-dot mask (also referred to as object substitution masking) has proven to be a valuable way to study the processes by which a visual object is consciously rendered (Goodhew et al., 2013). In this masking paradigm, a target is briefly flashed along with four surrounding dots, the latter of which persist for a variable duration. As this mask duration increases, target visibility is reduced, a phenomenon that likely reflects competition between the target and mask. Given the sparse nature of this mask, relative to more traditional masks that either fully surround the target's contours or spatially camouflage it, four-dot masking

is thought to be a powerful demonstration of interactions at a spatially global object level, rather than of spatially local contour interactions (Di Lollo et al., 2000). In addition to providing insights into the time-course of visual processing, four-dot masking offers a means to probe the way in which objects are individuated by the visual system (Goodhew et al., 2015; Lleras & Moore, 2003). It also allows us to examine how attentional manipulations can bias competition between visual objects, such as a target and mask (Pilling et al, 2014; Tata & Giaschi, 2004), and to explore the ability of a masked target to influence subsequent processing (Choo & Franconeri, 2010; Goodhew, Visser, Lipp, & Dux, 2011).

Until recently, four-dot masking was only reliably reported when the target was presented in parafoveal visual field, usually as part of a set of possible targets. Masking has also been found with a single target in parafoveal visual field, but only reported in cases where the spatial location of this target was randomized (Argyropoulos et al., 2013; Camp et al., 2015). A recent report, however, demonstrates four-dot masking involving a fully attended target in central visual field (Filmer et al., 2015). This important finding clearly shows that neither distributed attention nor crowding are necessary conditions for this form of masking. The ability to study four-dot masking in central visual field with a single target is valuable for a few reasons. With multiple targets, there is the risk of noise introduced by pooling data across multiple target locations. With a single target location, one can therefore obtain well controlled data with relatively fewer trials. Furthermore, the use of a single, central target allows for efficient use of space within the visual field, and there are thus fewer limitations on parameters such as target size, and target-mask separation. Finally, the use of a single target offers the theoretical convenience of removing the influence of crowding, and

feature misbinding (mistaking a feature of one of the distractors for a feature of the target), on any discovered effects.

In the current experiments, we sought to explore four-dot masking of a foveated target more closely. We were interested in two questions. First, how does a four-dot mask compare with an annulus mask, as a function of target-mask separation? As an annulus completely surrounds the target, it can mask the target through local inhibition as well as through object substitution/updating (Enns, 2004). It is unlikely, however, that a four-dot mask operates via local inhibition, given its sparse nature. Comparing these two mask types may provide evidence for a dissociation of these mechanisms. Second, we were interested in how this form of masking affects the target representation (Agaoglu et al., 2015; Harrison et al., 2016). To do this, we developed a matching task where observers adjusted the orientation of a bar to match that of the target bar, and examined how the distribution of errors changed between baseline and masking conditions. This task also allowed us to examine whether the errors were more likely to occur when the target bar was closest to any of the four dots[5]. If so, then this would point to the contribution of local masking mechanisms.

3.2 Experiment 3A

Our first step was to see whether we could successfully replicate the masking effect found in Filmer et al. (2015). In their study, they used a forward noise mask in addition to a common onset trailing four-dot mask, and found a reliable performance drop as the duration of the four-dot mask increased.

5 I would like to thank Haluk Öğmen for providing the suggestion for this experiment.

3.2.1 Stimuli

The target was an annulus with a bar projecting from its center to one of four cardinal points along the circumference. The inner radius of this annulus was 32.3 arcmin, and the outer radius was 39.5 arcmin. The projected bar, whose length was equal to the inner radius of the annulus, had a thickness of 5 arcmin. The mask comprised four small circles (diameter = 14.4 arcmin) located at the corners of an imaginary square concentric with the target. The separation of the mask, measured from the edge of each dot to the outer edge of the target annulus, was 31.6 arcmin. The forward mask was a circular patch of Gaussian noise, with a mean luminance equal to that of the background, and whose radius was equal to the outer radius of the target annulus. A square cross of length 28.7 arcmin was used for fixation. All stimuli were centrally presented with no spatial jitter. The target and four-dot mask are depicted in Figure 3.1.

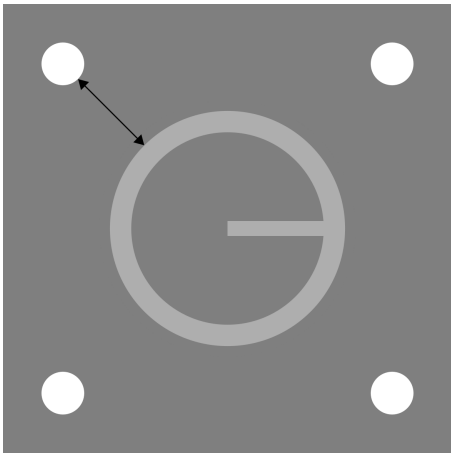

Figure 3.1. Target and four-dot mask drawn to scale. The black line indicates the distance along which target-mask separation was measured. The separation shown here was used in Experiment 3A.

The luminance of the target, as measured with a Konica-Minolta LS-100 (integration time: 400 ms) was 65.8 cd/m^2, while that of the four-dot mask and fixation cross was 97 cd/m^2. All stimuli were presented against a uniform background gray field of luminance 47 cd/m^2, on a VIEWPixx display calibrated to linear light (gamma = 1), at a viewing distance of 1.28 m in a dimly lit room. The display was running at 120 hz in scanning backlight mode. In this mode, the backlight is scanned down the display in synchrony with the updated pixels, and the pixel rise time (black to white) and fall time (white to black) are both 1 ms. Thus, in a single frame comprising a white field, each pixel lets light through (above and beyond the black level luminance) for only 2 ms. It is important to precisely specify the mode in which visual content is

rendered on the display, as this has implications for the actual stimulus durations, which are not always readily calculable based on reported frame rate (Elze, 2010).

3.2.2 Procedure

The trial sequence is shown in Figure 3.2A. Each run was initiated with a keypress, at which point the fixation cross appeared for one second. As soon as the fixation disappeared, the forward mask appeared, lasting 200 ms. Immediately following its offset were the target and four-dot mask (henceforth termed "4DM"). The target persisted for 8.3 ms. In the **common offset** condition, which served as a baseline, the mask disappeared with the target (mask duration = 8.3 ms). Two other mask durations were tested: **250 ms**, and **500 ms**. Upon mask offset, the background remained visible, and a keypress response indicated which of the four directions the observer believed the line to be pointing along (up, down, left, or right). The task was self paced, and keypress responses initiated the next trial. Within each run, each of the three mask durations was repeated 30 times in random order, and observers completed three runs, for a total of 90 trials per condition. Seven observers completed this experiment, one of whom is the first author, and the rest of whom were naive. All observers, in this and subsequent experiments, gave verbal consent before participation. The experiments in this chapter were approved by the Human Participants Review Committee (HPRC) at York University (approval number HPRC 2014-094). Before beginning these experimental runs, each observer completed a PEST procedure (Taylor & Creelman, 1967) for the common offset condition, where the standard deviation of the noise patch was varied until 80 percent performance was achieved. This value was then used in the main experiment for that observer across all conditions. Feedback was not provided in either the PEST phase or the experimental trials.

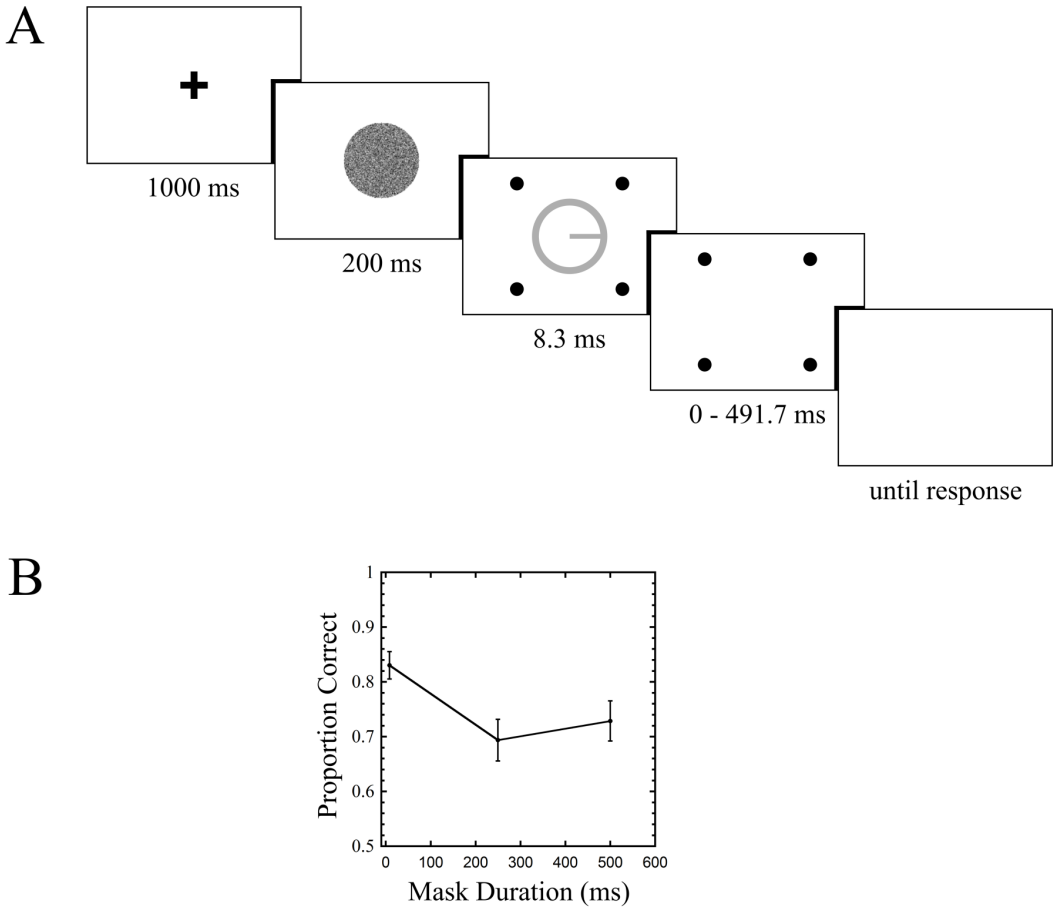

Figure 3.2. Trial sequence and results for Experiment 3A. Stimuli here and in all subsequent figures are rendered in reverse contrast. Error bars indicate standard errors.

3.2.3 Results & Discussion

Results are shown in Figure 3.2B. A one way repeated measures ANOVA revealed a main effect of mask duration ($F(2,12) = 4.613$, $p = 0.033$). Follow up comparisons showed a difference between the common offset and 250 ms mask duration conditions ($t(6) = 3.985$, $p = 0.007$), no difference between common offset and 500 ms

($t(6) = 1.942$, $p = 0.1$), and no difference between 250 ms and 500 ms ($t(6) = 0.681$, $p = 0.521$).

Experiment 3A shows a clear masking effect, with about a 13.5 percent drop in performance between the common offset and 250 ms mask duration conditions. The data do not provide strong evidence for masking at the 500 ms condition, nor do they support recovery from masking (Goodhew, Dux, Lipp, & Visser, 2012), as there was no difference between the 250 ms and 500 ms conditions. In our next experiment, we measured masking as a function of target-mask separation, with both a 4DM and an annulus mask.

3.3 Experiment 3B

Our next experiment compared a 4DM with an annulus mask as a function of the separation between target and mask. If the annulus mask is acting primarily through local inhibitory mechanisms, while the 4DM involves object-level mechanisms, then as target-mask separation increases, one would expect a drop in annulus masking, while dot masking should remain relatively unchanged.

3.3.1 Procedure

The general procedure was similar to the previous experiment. Here, runs were blocked according to mask type (4DM, annulus, see Figure 3.3). Within each run, four different mask separations were tested, with a constant mask duration of 250 ms (**7.2 arcmin, 17.2 arcmin, 27.3 arcmin,** and **103.5 arcmin**), in addition to a common offset condition for each mask at the lowest separation (7.2 arcmin). Each of these five conditions was tested in random order, 30 times per run, and observers completed three runs for each mask type, for a total of 90 trials per mask type per separation. Nine observers completed this experiment, all of whom were naive. Observers completed runs

in alternating order of mask type. Four of them began with the annulus mask, and five began with the 4DM. Instead of running a PEST procedure, observers were trained in the common offset condition, with increasing amounts of noise in the forward mask. This was done until the experimenter had established a noise level at which performance was around 75 percent correct. In these training trials, a correct response generated an auditory tone. In the experimental trials, in this and in all other experiments in this chapter, no feedback was provided.

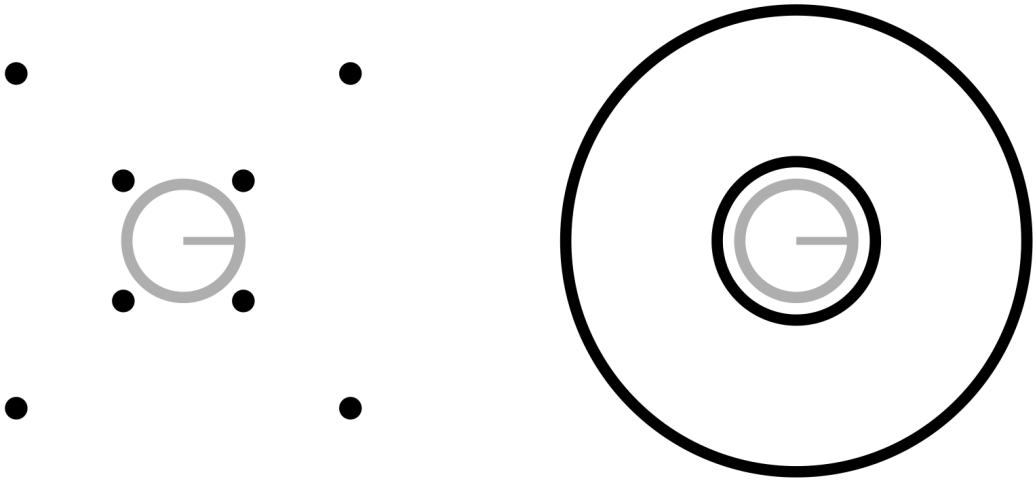

Figure 3.3. Stimuli used in Experiment 3B. **Left:** Four-dot mask. **Right:** Annulus mask. Each mask type is shown with both the smallest (7.2 arcmin) and largest (103.5 arcmin) target-mask separations. Stimuli are drawn to scale.

3.3.2 Results & Discussion

Results are shown in Figure 3.4. A two way repeated measures ANOVA (first factor = mask type: two levels (4DM, annulus); second factor = target-mask separation: four levels, excluding baseline) revealed no main effect of mask type ($F(1,8) = 1.218$, $p = 0.3$), a main effect of separation ($F(3,24) = 7.489$, $p = 0.001$, and an interaction between mask type and separation ($F(3,24) = 3.245$, $p = 0.04$). To verify that masking was obtained with the 4DM, a paired sample t-test was used to compare the 7.2 arcmin 250 ms mask with the baseline condition (7.2 arcmin common offset mask), and this confirmed masking ($t(8) = 2.861$, $p = 0.021$). The striking pattern found with the annulus mask is strong evidence of the primacy of spatially local inhibitory mechanisms here. By a separation of 27.3 arcmin, annulus masking was greatly attenuated, compared to the 7.2 arcmin mask ($t(8) = 3.067$, $p = 0.015$). In contrast, the 4DM showed no change in masking between the 7.2 and 27.3 arcmin conditions ($t(8) = 0.77$, $p = 0.47$).

While other studies have looked at the effect of separating the mask, *as a whole,* from the target (e.g. Jiang & Chun, 2001), to our knowledge, only one other study has looked at the effect of increasing the separation of the individual dots, as was done in our current experiment (Di Lollo et al., 2000). In that study, masking did not decrease, up to the measured separation of 40 arcmin, a finding that closely matches our own. These findings support the idea that (spatially sensitive) local inhibitory mechanisms underlie annulus masking, while (less spatially sensitive) object level mechanisms underlie dot masking.

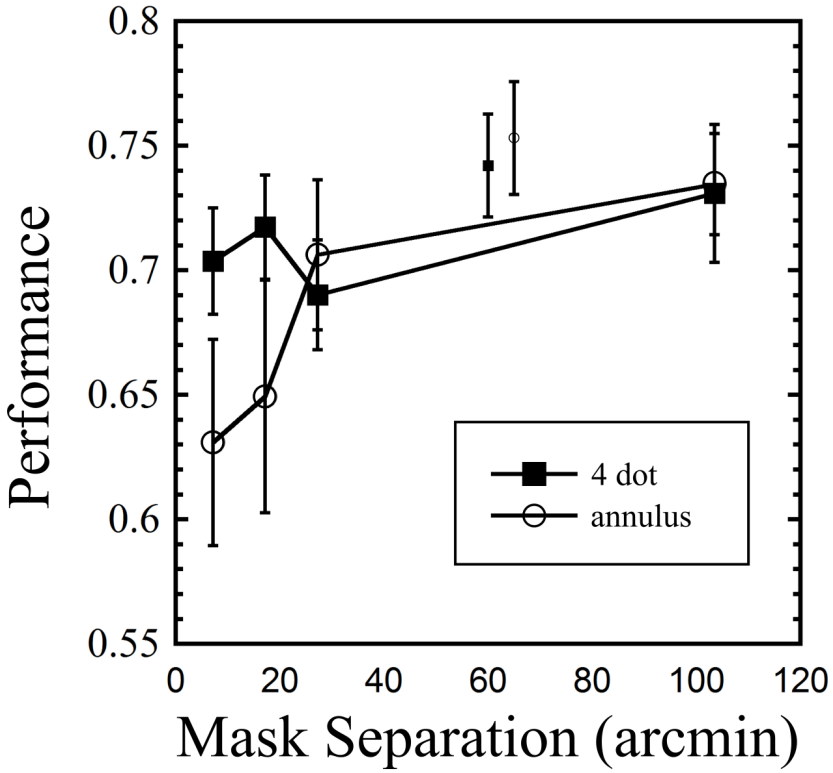

Figure 3.4. Results from Experiment 3B, showing performance across nine observers for both mask types as a function of separation from target. The two smaller symbols in the upper middle region indicate baseline values for the two masks (common offset mask, 7.2 arcmin separation). Error bars indicate standard errors.

3.4 Experiment 3C

The two previous experiments had observers select from four possible line orientations. In our final experiment, which only used a 4DM, the task was to adjust the orientation of a reference line until it matched one of 90 possible target orientations. This allowed us to investigate two separate questions. First, are errors more likely to occur when the target line is oriented such that it is closer to one of the four dots? If so,

then this would point towards the contribution of spatially local masking mechanisms. Second, how does the distribution of errors change between baseline and masked conditions? Is the representation of a masked target simply less precise (in the orientation domain), or is it never consciously processed, in which case observers would have to guess? Or is it some combination of both? By modelling the errors as a mixture of a Gaussian distribution centered on the actual target orientation (representing trials in which the target was perceived) and a Uniform distribution across all possible orientations (representing trials in which observers were relegated to guessing), we were able to address this question.

3.4.1 Procedure

The stimuli were similar to the previous experiments; however, the target could now adopt any of 90 unique orientations, spaced at 4° intervals around the circle, ranging from 0° to 356°. Observers used the mouse to adjust and select the orientation of a reference pattern, which appeared 500 ms after target offset. The reference pattern and the target were visually identical, except that the orientation of the target was determined by the current trial, whereas the orientation of the reference pattern was determined by the current mouse position.

Each run contained three conditions. **Target alone** with no mask, a **common offset** 4DM, and a **250 ms** duration 4DM. The target-mask separation was 10.8 arcmin. Within each run, each condition was tested once at each of the 90 orientations, for a total of 270 trials per run, which were presented in random order. Within each run, a message appeared every 90 trials, indicating the current progress, at which point a mouse click returned the observer to the trial flow. Each observer completed three runs, thus each orientation was tested three times per condition. Nine observers

completed this experiment, one of whom is the first author and the rest of whom were naive. Training was done in the common offset condition, and for each observer, the experimenter increased the standard deviation of the noise patch until performance was about 60 percent correct. In this training phase, a correct response was defined as being within 20 degrees of the actual target orientation, and was signalled with a tone.

3.4.2 Results & Discussion

The purpose of this experiment was two-fold. First, we investigate whether there was a bias for increased masking when the orientation of the target line was closer to the four dots. Second, we model the distribution of errors for each of the three masking conditions.

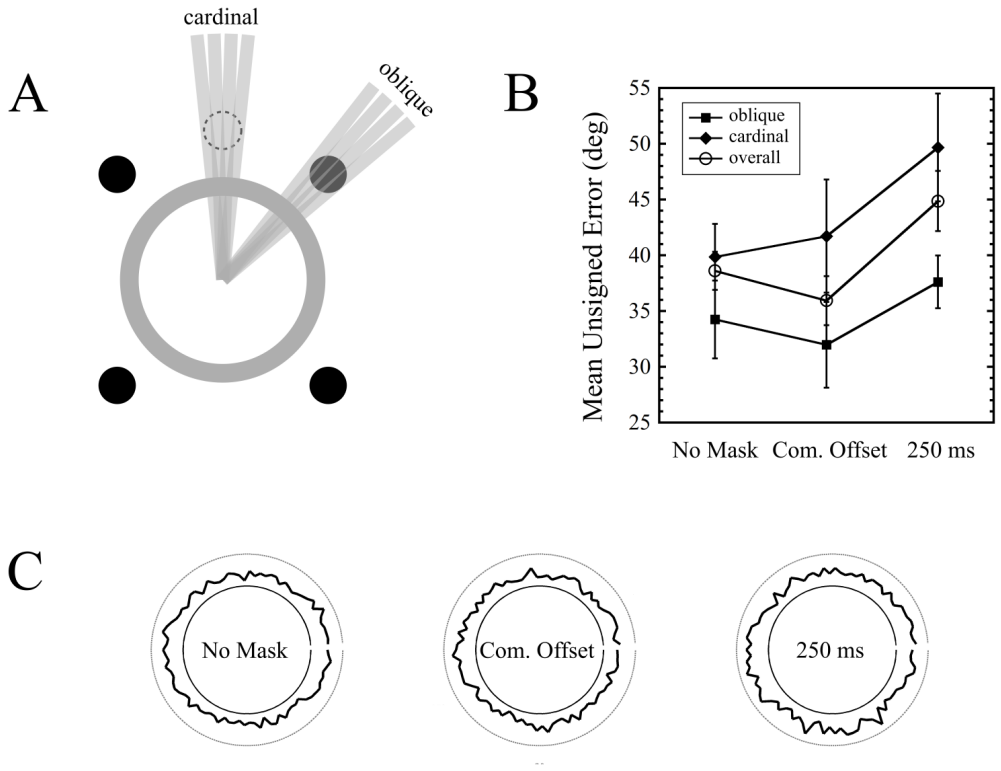

Figure 3.5. Orientation bias results from Experiment 3C. **A:** Depiction of how oblique and cardinal orientations were defined. The dot shown with the dashed outline served as a marker to classify orientations as oblique (only one of these dots is shown in the figure). **B:** Errors as a function of mask condition and orientation type (oblique vs cardinal) across nine observers. Error bars indicate standard errors. **C:** Polar plots depicting errors as a function of target orientation, for each mask type, across all observers. The magnitude of the error is depicted by the radially projected distance from the inner circle to the thick black curve. The outer circle indicates the error expected by chance performance (90 degrees).

For each trial, the difference between the actual target orientation and the reported orientation was calculated. Trials were pooled into two categories: those in which the presented target line orientation was aligned in proximity to the dots

(**oblique**), and those in which it was aligned along the cardinal axes (**cardinal**), and thus maximally distant from the dots. Classification was done by visual inspection of a modified target pattern, which was identical to the experimental pattern, except that the oriented bar now extended beyond the annulus (Figure 3.5A). Upon inspection of each of the 90 bar orientations, if any part of the bar intersected any of the four dots, the orientation was classified as oblique. Cardinal orientations were classified with a similar procedure, but with the four dots arranged in a diamond pattern (see Figure 3.5A for a depiction of one of these dots). Of the set of 90 orientations, 18 orientations were classified as oblique, and 18 as cardinal. This meant that for each observer, there were 54 oblique trials and 54 cardinal trials for each of the three masking conditions. Results are shown in Figure 3.5B. Note that the errors are unsigned. The overall errors indicate the mean of all 90 orientations.

The first thing to note is that masking did occur. The overall error for the common offset condition was 35.9 degrees, while that of the 250 ms mask was 44.9 degrees, and this difference was significant ($t(8) = 4.122$, $p = 0.003$). Interestingly, the presence of a common offset mask increased performance relative to no mask at all ($t(8) = 3.822$, $p = 0.005$).

If the (trailing) 4DM was exerting its effect through local spatial mechanisms, then a greater increase of errors in the oblique orientations in the 250 ms condition (relative to the common offset baseline) would be expected, compared to the increase in cardinal errors. To test this interaction, we ran a two way repeated measures ANOVA (first factor = mask condition: two levels (common offset, 250 ms); second factor = error location: two levels (oblique, cardinal)). This revealed a strong trend for the effect of mask condition ($F(1,8) = 4.825$, $p = 0.059$; no main effect of error location ($F(1,8) = 3.549$, $p = 0.096$; and no interaction ($F(1,8) = 0.138$, $p = 0.72$).

Our data do not show evidence of an orientation bias in masking: we did not observe increased masking at target orientations proximal to the dot locations, as evidenced by the almost perfectly parallel lines in the right half of Figure 3.5B. While this does not definitively prove that dot masking operates via spatially global mechanisms, it is consistent with the finding in Experiment 3B that dot masking is, to a large extent, independent of target-mask separation. The finding that across all masking conditions (target alone, common offset, and 250 ms mask), observers did not perform worse on oblique orientations is puzzling (in fact, there was a trend towards better performance on oblique orientations), given the well documented oblique effect, where orientation judgments are poorer for oblique, relative to cardinal orientations (Jastrow, 1892; Heeley & Timney, 1988). One possible explanation for this trend of an inverse oblique effect we found in our data may be that the dots in the four dot mask, which were aligned with the oblique orientations, aided in perception of these lines, perhaps through collinear facilitation (Polat & Sagi, 1993). A way to test this would be to repeat Experiment 3C but with the four dots arranged in a diamond configuration. If collinear facilitation plays a role here, then the inverse oblique effect may be reversed. Another possibility is that information was pooled over a larger area in the oblique orientations, and this area was large enough to incorporate the dots. Such an explanation has been offered for an inverse oblique effect found with translational glass patterns (Wilson, Loffler, Wilkinson, & Thistlethwaite, 2000), and indeed, the degraded nature of our target stimulus may have allowed such a pooling effect to manifest even without the masking dots. We ran a simple control experiment to investigate the relationship between dot position and performance, where a 250 ms four dot mask was presented in either a square configuration (as in the previous three experiments), or in a diamond configuration, where the dots were rotated 45 degrees relative to the square position. As

in Experiments 3A & B, the target could be oriented in one of the four cardinal orientations. If dot-target alignment facilitates performance, then observers would perform better in the diamond configuration, since here the four possible target orientations align with masking dots. Seven observers completed 90 trials with each of the mask types, and no difference in performance was found ($t(6) = 0.3$, $p = 0.774$). This suggests that the trend of an inverse oblique effect found in Experiment 3C was likely not due to the orientation of the dots. Either way, for the present purpose, what is important here is the difference in performance between the common offset and 250 ms masking conditions in Experiment 3C, and this difference in performance was found to be equivalent across both sets of orientations.

3.4.3 Mixture Model Analysis

For any trial in which masking occurred, there are at least two possible scenarios. On the one hand, the target may have never been consciously processed, either because the representation was completely obliterated, or because it had been degraded below a critical threshold. In such a scenario, an observer would have to guess the orientation of the target. On the other hand, the target representation may have remained consciously accessible, but in a degraded form. If the nature of this impoverishment was such that the precision of orientation information was compromised, then we would expect a reduction in the precision of the response. Note that these possibilities are not mutually exclusive. We fit a mixture model to our data to estimate the relative proportions of trials that corresponded to these two possibilities, using Gaussian and Uniform distributions for the perceived target and non perceived target trials, respectively.

The probability density function for this mixture model, which defines the distribution of signed errors, is defined below in Equation 3.1:

$$PDF = W_G * G(\mu, \sigma) + (1 - W_G) * U(-\pi, \pi) \quad \textbf{Equation 3.1}$$

where G is a Gaussian function, with mean (μ) and standard deviation (σ), and U is a Uniform distribution along the specified interval. W_G is the weight of the Gaussian term, and as the mixed distribution sums to unity, the weight of the Uniform term is 1-W_G. In the context of our experiment, a lower value of W_G indicates more guessing, and a higher value of σ indicates less precise responses of perceived targets. This mixture model is very similar to that used by Zhang & Luck (2008), except they used a circular normal distribution, whereas we used a normal distribution. While a circular normal distribution would be a better match for the circular nature of our error variable, which spanned -180° to 180°, it has been shown that a single Gaussian is an excellent approximation to a circular normal distribution when the standard deviation is small relative to the range of possible errors (Shooner, Tripathy, Bedell, & Öğmen, 2010; also see Appendix A in Agaoglu et al., 2015).

For each observer, and for each of the three masking conditions (no mask, common offset, 250 ms mask), a maximum likelihood procedure (Myung, 2003) was used to estimate W_G, μ , and σ. The results, pooled across all nine observers, are shown in Figure 3.6. A one way repeated measures ANOVA was run on each of the two parameters, σ and W_G, across all three masking conditions. For σ, there was no main effect of masking condition ($F(2,16) = 0.046$, $p = 0.956$). For W_G, a main effect of masking condition was found ($F(2,16) = 12.215$, $p = 0.001$). Interestingly, while the guess rate with the trailing mask was greater than in the common offset condition (t(8)

= 4.664, p = 0.002), it was also higher in the no mask condition than in the common offset condition (t(8) = 3.985, p = 0.004).

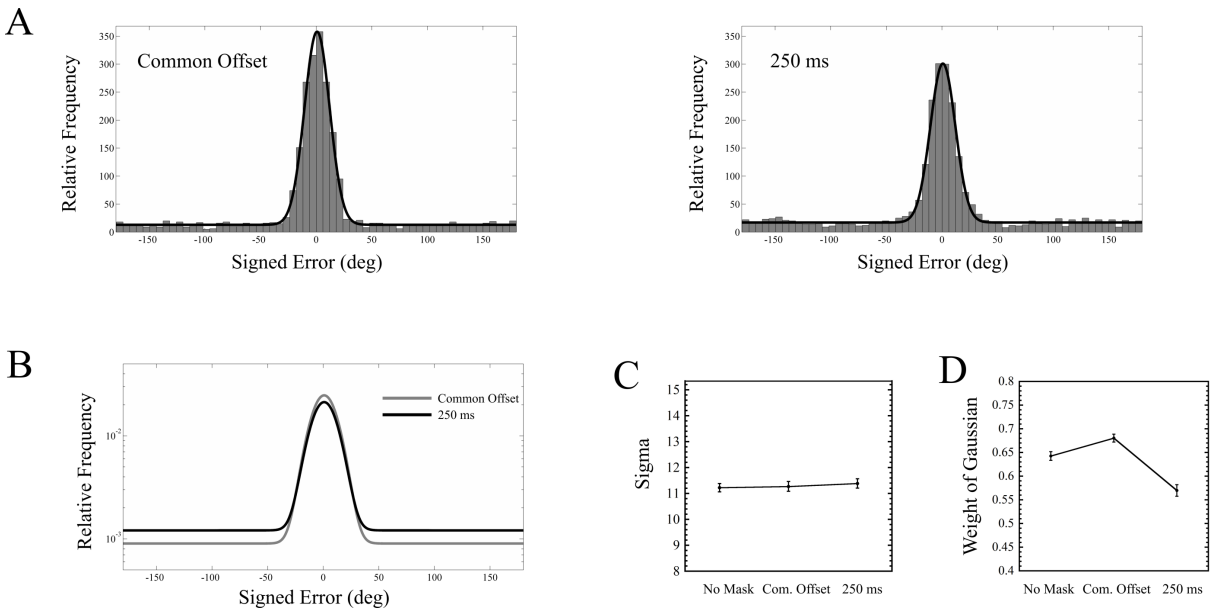

Figure 3.6. A: Distribution of errors across nine observers for the common offset and 250 ms conditions, along with model fits. **B:** Model fits for the error distributions of the common offset mask and 250 ms mask (note the log scale). Each curve represents a mixture of a Gaussian and Uniform. Note that the width of the Gaussian remains unchanged between these two masking conditions. **C:** Standard deviation of Gaussian component for each of the three mask conditions. **D:** Weight of Gaussian component for each of the conditions. Error bars indicate standard errors.

Masking appears to have absolutely no effect upon the precision of responses (Figure 3.6C). Instead, the performance drop between the common offset mask and the trailing mask (Figure 3.5B) is driven entirely by the target being rendered invisible, where the guessing rate increased from about 32% to 43% (Figure 3.6D). This also demonstrates that the 250 ms mask had a measurable impact, relative to the common offset mask, on only a fraction of the trials.

These results are fairly striking, especially when compared against two recent studies that conducted very similar experiments (Agaoglu et al., 2015; Harrison et al., 2016). In particular, Agaoglu et al. used an oriented bar as a target stimulus very similar to ours, and, using a backward masking paradigm, found that the standard deviation of observers' responses increased from about 11 degrees to as high as 20 degrees, depending upon the type of mask being used (all mask types also increased the guessing rate). On the other hand, the standard deviation of the responses in our experiment remained at around 11 degrees across all masking conditions. Harrison et al. (2016), using a similar masking paradigm, where observers attempted to match the orientation of a gap in a target annulus, found that the standard deviation of responses showed about a 40 percent increase in the trailing vs. common offset condition. We will explore this discrepancy further in the next section.

3.5 General Discussion

In Experiment 3A, we successfully reproduced four-dot masking in a fully attended and foveated target, as originally reported by Filmer et al. (2015). The data from Experiment 3B, which directly compared a 4DM with an annulus mask, strongly suggest that these two mask types operate through at least partially distinct mechanisms. Experiment 3C demonstrates that in the context of our stimuli and task, there is no evidence that the encoding precision of the target (in the orientation domain) is affected by masking. Rather, masking appears to be driven exclusively by rendering the target inaccessible to conscious processing. In this section, we will explore these findings in more depth.

As in Filmer et al. (2015), we included a forward noise mask in our experiments. It is an open question, however, whether similar results would obtain had we

incorporated the noise mask into the same frame as the target (although see discussion of this point in Chapter 5.4). In the current experiments, while the stimulus onset asynchrony (SOA) of the noise mask, relative to the target, was -200 ms, the onset of the target occurred immediately following the offset of the noise mask (i.e. if we take into account the scanning mode of our display, this would correspond to an interstimulus interval (ISI) of slightly below 8.3 ms). Given this small delay between the offset of the noise mask and the onset of the target, crosstalk between successive frames may have allowed information from both stimuli to be present on the same frame, and therefore integrated into a unified percept. Indeed, while the VIEWPixx display advertises a pixel fall time of 1 ms, a recent study found measurable crosstalk in this display (Baker, Kaestner, & Gouws, 2016), suggesting that the advertised pixel fall time may have been based on equipment not sensitive enough to capture the low luminance "afterglow". Alternatively, retinal aftereffects produced by the relatively high contrast noise patch may have persisted long enough to produce a negative afterimage of the noise patch which then integrated with the relatively low contrast target. In either case, the stimuli certainly appeared to be integrated when presented on the display. Inspection of neurophysiological data in Macknik & Livingstone (1998) suggests that the transient response to a target is slightly lower with an SOA of -100 ms (ISI of 0 ms), relative to a common onset common offset condition, (Figure 4 in their paper). However, the masks in that study comprised a set of flanking bars. A more relevant study (Agaoglu et al., 2015), which used an almost identical task to ours, showed peak masking with a noise mask at an SOA of 0 ms, which appeared to show even greater masking than that found with a forward mask of SOA -10 ms (ISI = 0 ms). Thus, although we have not tested this, we suspect that the principal element that enables

successful four-dot masking here is the incorporation of noise into the target, rather than the use of a forward mask.

Why would a noise mask be critical for four-dot masking? Its ostensible purpose in Filmer et al. (2015) was to reduce baseline performance to below ceiling. Pilot work in our lab, however, showed no evidence of four-dot masking (in the absence of a noise patch) when the contrast of the target was reduced such that baseline performance was well below ceiling. Indeed, if ceiling effects were the only issue here, then it is surprising that it has taken almost two decades to produce reliable four-dot masking in a fully attended and foveated target (although see Dux, Visser, Goodhew, & Lipp, 2010). Rather, it is more likely that the target needs to be degraded in a specific manner. With the use of a noise mask, the integrity of the stimulus as a coherent object is dramatically compromised. In order to perceive an oriented bar, perceptual filling in processes may be required, and this may delay or otherwise compromise the formation of a stable object representation, which in turn may leave the target more vulnerable to being dominated by the mask. Another possibility is that if the 4DM is able to disrupt the perception of target elements proximal to the dots, as is the case, for example, in *object trimming* (Kahan & Enns, 2010), this may sometimes leave less information available for the visual system to glean the bar's orientation, if the part of the object being trimmed happens to contain useful information. This would be equivalent to reducing the signal to noise ratio. Both of these possibilities are compatible with the finding in Experiment 3C that suggests when an object is successfully masked here, it is rendered invisible.

It should be noted that while Experiment 3A was very similar to that in Filmer et al. (2015), it was not an exact replication, as there were differences in the dimensions and contrast of the stimuli. In particular, while the target and (four dot) mask in

Filmer et al.'s first experiment had the same contrast relative to the background, we used a lower contrast for the target, compared to the (four dot) mask. Our stimuli are thus more comparable to Filmer et al's second experiment, where the contrast of the target was thresholded to the desired performance level.

The data from our second experiment provides evidence for a number of related ideas. First, the finding that dot masking did not attenuate with increasing target-mask separation suggests that four-dot masking operates via mechanisms that are relatively insensitive to spatially local interactions between target and mask. Rather, masking here likely involves mechanisms that primarily involve representations at higher levels of the visual processing hierarchy, where receptive fields are larger (Dumoulin & Wandell, 2008; Smith, Singh, Williams, & Greenlee, 2001). While the involvement of higher visual areas is compatible with object level accounts involving interference between feedback and ongoing input, our data do not directly say anything about whether feedback is involved. Second, the finding that annulus masking *did* attenuate with increasing target-mask separation is strong evidence that here, the mechanisms are sensitive to spatially local interactions between the target and the mask, and confirms the findings of many previous studies that have found a similar sensitivity to spatial separation in the literature on metacontrast masking (see § 2.6.6 in Breitmeyer & Öğmen, 2006). Finally, the interaction between mask type and separation adds to the evidence that four-dot masking operates through mechanisms that are at least partially distinct from those involved with masks whose contours fully surround the target.

In our final experiment, we found that masking appears to be driven exclusively by rendering the target completely invisible. While this finding is certainly compatible with the idea of object updating (Pilling & Gellatly, 2010), where the original object (target + mask) is updated into a new object (mask), it is not proof of object updating.

Another possibility is that four-dot masking simply degrades the target without impacting the encoding precision of its orientation information. For example, if the mask is simply reducing the contrast of the target, then it is possible that orientation encoding is unaffected. Mareschal & Shapley (2004) found that orientation discrimination of foveally presented gratings was unaffected by contrast when the diameter of these gratings was large (1 degree of visual angle). However, for two out of the three observers in their experiment, contrast did have a dramatic impact on these thresholds when the diameter was as large as 0.5 degrees (all of the observers showed reduced performance as contrast was reduced when the diameter was 0.25 degrees). The length of the oriented bar in our experiment was about 0.5 degrees, and it is certainly possible that our cohort of observers was demonstrating contrast invariance to orientation encoding with this stimulus size. It would be interesting to run our experiment with a bar length of 0.25 degrees and see if precision is similarly unaffected. Indeed, in Filmer et al. (2015), the diameter of the target was 0.55 degrees, which meant that the bar length was about 0.25 degrees. Another alternative to object updating is object trimming (Kahan & Enns, 2010), where a two dot mask was shown to interfere with target contours that were adjacent to these dots. If object trimming was occurring in our experiment, then, as discussed earlier, this could account for the target being rendered invisible due to a reduced signal to noise ratio. However, if this were the case, then we might expect more masking when the target orientation was aligned with the oblique axes, and we found no evidence of this (Figure 3.5B).

Another aspect of Experiment 3C worth considering is that, while we found no change in response precision, two other similar studies showed a substantial drop in response precision (Agaoglu et al., 2015; Harrison et al., 2016). However, there are some important differences to consider here. In Harrison et al., the target was an annulus

whose gap could adopt a number of different orientations. It is possible that the orientation specific information in such a stimulus may be compromised, while that of an oriented bar would not, given the same general type of mask induced degradation. For example, it is clear how object trimming could result in part of the circle being occluded near the gap, resulting in a larger perceived gap. In such a situation, response precision could certainly suffer. In Agaoglu et al., while the target was very similar to ours (although their oriented bar was almost twice as long as ours), they did not use a four-dot mask. Rather, they compared masking functions between an annulus mask, a noise mask, and a structure mask (the last of which comprised three bars similar to the target bar but in random orientations). It is easy to conceive how, for example, their structure mask could add noise in the orientation encoding domain. Furthermore, the parameter in these masking functions was the SOA of a pulsed mask, rather than the mask duration of a common onset mask. Finally, in both these studies, stimuli were presented non foveally. In Harrison et al., the target was one among anywhere from two to eight possible targets presented at 3.5 degrees from fixation. In Agaoglu et al., a single target was presented at a 6 degree horizontal eccentricity. While orientation discrimination thresholds do rise as a function of eccentricity (Sally & Gurnsey, 2004), this alone does not explain why precision would be reduced with the addition of a para or metacontrast mask, as shown in Agaoglu et al., or with an increased four-dot mask duration, as in Harrison et al. However, there may be an interaction between contrast and eccentricity. In Mareschal & Shapley (2004), the effect of contrast upon orientation discrimination thresholds was evident at larger target sizes in the parafovea compared to the fovea. When stimuli were presented at 5 degrees of horizontal eccentricity, there was a large effect of contrast upon orientation discrimination of gratings that were as large as 1 degree, whereas in the fovea, all observers showed contrast invariance at this

stimulus size. In Agaoglu et al., the oriented bars were 1 degree in length, and were presented at 6 degrees of eccentricity. If their masks were operating through contrast reduction, then this is a plausible explanation of the reduced response precision.

The current set of experiments reinforces Filmer et al.'s (2015) finding that four-dot masking can be reliably obtained with a fully attended and centrally presented target, an idea which must be accounted for in any theory of masking via object updating. The direct comparison between the two mask types used in Experiment 3B adds new evidence that four-dot masking operates via unique mechanisms when compared to traditional masking stimuli. Our last experiment provides novel insights into the way in which four-dot masking impacts the visual representation of a foveated target.

On a final note, we have chosen to refrain from making assumptions about the underlying mechanisms of masking, beyond those that can be reasonably supported by our data, and this is reflected in our terminology throughout this chapter. While our findings are consistent with an object mediated account of masking, we have not introduced any experimental manipulations that test, for example, whether object *updating* (versus object *substitution*) accounts for our masking (Enns et al., 2010). Accordingly, we have avoided the use of the terms object substitution and object updating wherever possible. Similarly, as our data say nothing about the relationship between feedforward and feedback signals (Kafaligonul et al., 2015), we have avoided implicitly assuming any particular processing regime.

Chapter Four

Computational Model

SUMMARY

To assess our data in the light of understood properties of the visual system, we developed a computational model and compared its output against our data across both studies. The model can be summarized as follows. Mask and target stimuli are first spatially and temporally filtered, and then transformed using magnocellular (M) and parvocellular (P) contrast gain functions. A visual response function (VRF) is then obtained by first dividing the P representation of the target by the M representation of the mask, and then spatially averaging the result. During this step, the target's P representation is also divided by a spatially uniform component to simulate masking at the object level. Finally, the VRF is averaged over time to produce a single visual response value for a given masking trial.

4.1 Model Description

For each trial, a single target and a single mask are each represented as a 600x600 pixel grayscale image $\mathbf{L(x,y)}$, where $\mathbf{L}$ is the unnormalized luminance across the x and y dimensions. The stimuli in our experiments were presented on a gray background whose luminance was half that of the peak display luminance. Thus, the contrast of any given pixel could be increased by either increasing or decreasing its luminance relative to the background luminance. In the model, each image is transformed from luminance values into contrast values $\mathbf{C(x,y)}$, where the contrast of each pixel is the difference between the pixel's luminance and the background luminance

(L_B), divided by the pixel's luminance plus the background luminance (see **Eq. 1**). This is somewhat analogous to the Michelson contrast of an image.

$$C(x,y)=\frac{|L(x,y)-L_B|}{L(x,y)+L_B} \qquad \textbf{Equation 4.1}$$

The reason for transforming luminance into contrast is twofold. First, the human visual system is more sensitive to contrast than it is absolute luminance. Second, it is important in our model that the background pixels of our stimuli do not contribute towards masking. In the contrast domain, these background pixels have a value of 0, whereas in the luminance domain they would have a positive value.

Next, these contrast images are spatially filtered by the appropriate receptive field functions, **RF(x,y)**, giving **F(x,y,t)**.

$$RF_M(x,y)=1.8\cdot e^{-\left(\frac{x^2+y^2}{0.128^2}\right)}-0.8\cdot e^{-\left(\frac{x^2+y^2}{0.192^2}\right)} \qquad \textbf{Equation 4.2a}$$

$$RF_P(x,y)=1.8\cdot e^{-\left(\frac{x^2+y^2}{0.032^2}\right)}-0.8\cdot e^{-\left(\frac{x^2+y^2}{0.048^2}\right)} \qquad \textbf{Equation 4.2b}$$

Here, and in all other equations, the subscripts M and P refer to magnocellular and parvocellular components, respectively.

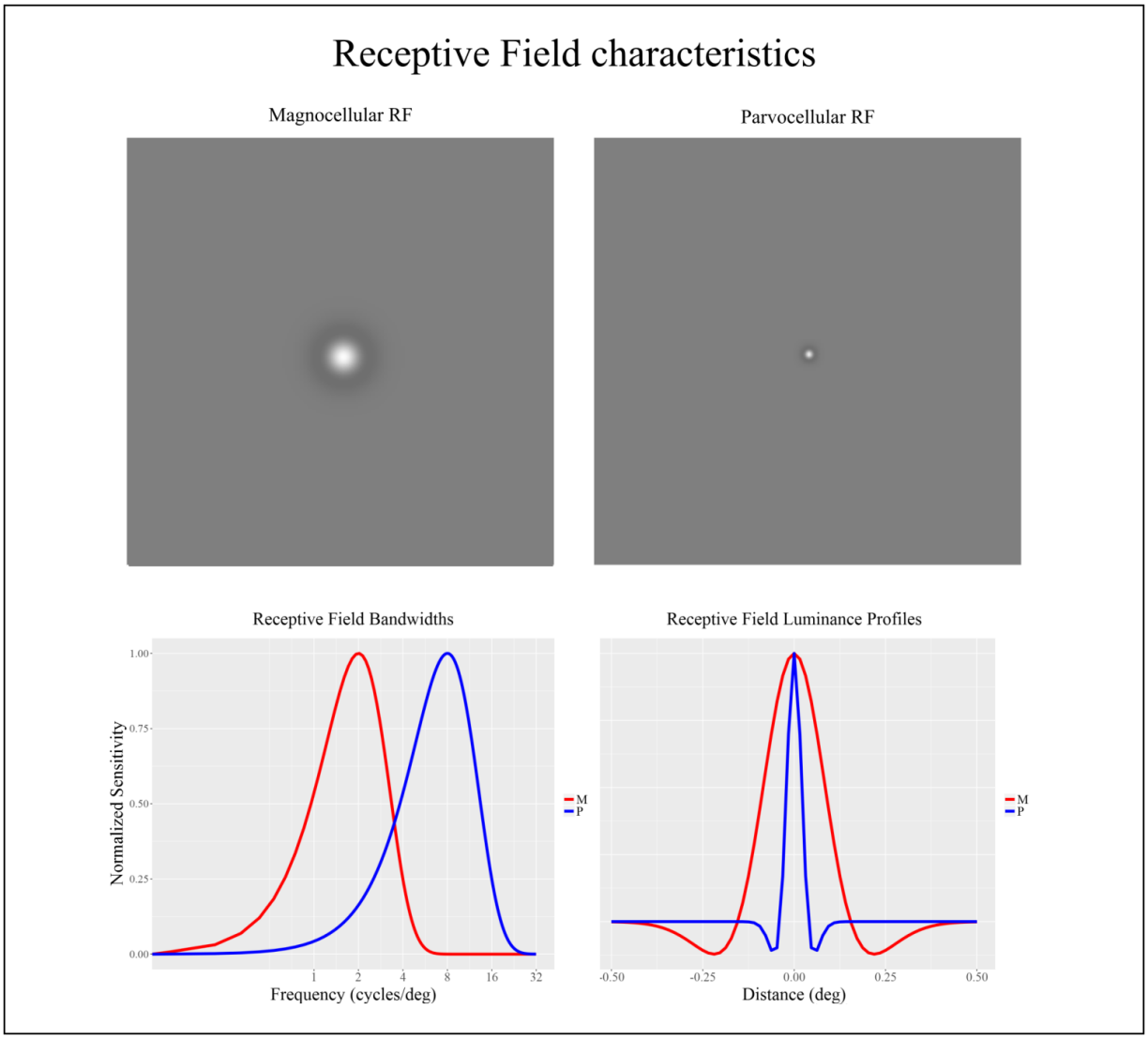

Figure 4.1. Magnocellular and Parvocellular receptive fields. The **top** panels show two dimensional plots of Equation 2. To allow a visualization of the negative regions, the values have been scaled and shifted so that a value of 0 is now represented by mid gray (background). For clarity, the visible region has been cropped from the original 600x600 pixel image, and is magnified by a factor of 2.7. The **bottom left** panel shows the Fourier transform of these functions, and illustrates the peak spatial frequencies of these receptive field filters (red = magno, blue = parvo). In the **bottom right** are shown luminance profiles across space for both filters. These last plots were obtained by profiling the two dimensional plots through their centers.

The M and P filters chosen for our model are difference of Gaussians, with peak spatial frequencies of 2 and 8 cycles per degree, for M and P, respectively. In Equation 2, the space constants are in units of degrees of visual angle. In the model, these were appropriately scaled into the pixel domain. The filtered images are normalized based upon the filter response to gratings of optimal spatial frequency.

Next, the timecourses for the target and mask stimuli, $\mathbf{T(t)}$, are convolved with M and P temporal impulse response functions, $\mathbf{I(t)}$, and then halfwave rectified. The timecourses represent, with a precision of 1 ms, the periods during which a stimulus is present and absent, and are encoded as a one dimensional array of ones (present) and zeros (absent). The length of the timecourse for all trials was set to 300 ms. The M and P impulse responses (Equation 4.3) are based on recordings of cells in the macaque striate cortex (De Valois & Cottaris, 1998). As with the spatial filtering, the convolved responses are normalized based upon the maximum response for each channel (see description in Figure 4.3).

$$I_M(t) = \left\| \left(\frac{t}{5}\right)^{15} \cdot e^{\left(\frac{-t}{5}\right)} - 0.8 \cdot \left(\frac{t}{6}\right)^{15} \cdot e^{(-t/6)} \right\| \qquad \textbf{Equation 4.3a}$$

$$I_P(t) = \left\| \left(\frac{t}{6}\right)^{15.26} \cdot e^{\left(\frac{-t}{6}\right)} - 0.1 \cdot \left(\frac{t}{8}\right)^{15.26} \cdot e^{(-t/8)} \right\| \qquad \textbf{Equation 4.3b}$$

$$N_M(t) = [I_M(t) * T(t)]^+ \qquad \textbf{Equation 4.4a}$$

$$N_P(t) = [I_P(t) * T(t)]^+ \qquad \textbf{Equation 4.4b}$$

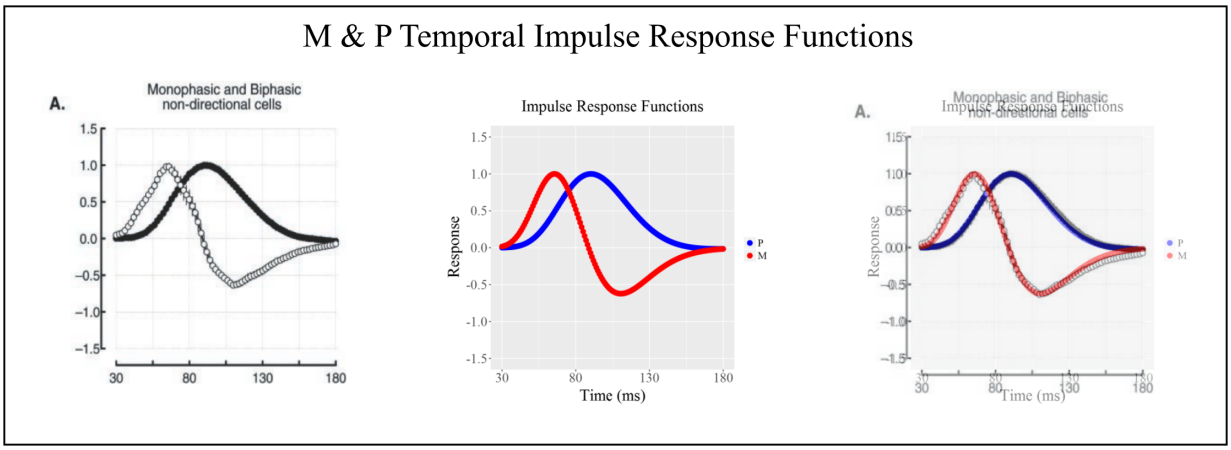

Figure 4.2. Magno and parvocellular temporal impulse response functions. **Left:** electrophysiological data from De Valois & Cottaris (1998). **Middle:** functions used in our model (Equation 4.3). Note that these are normalized to have a peak response of 1. **Right:** Electrophysiological data superimposed over the functions we used in our model.

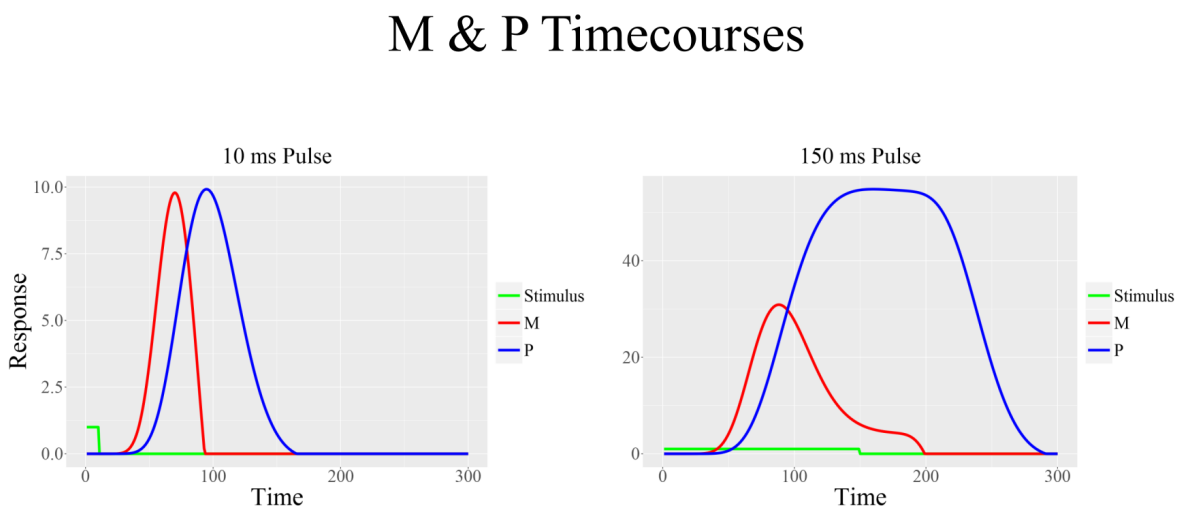

Figure 4.3. Magno and parvocellular response functions (N(t)), along with stimulus timecourses shown in green (T(t)). Shown here are unnormalized results of two stimulus schedule convolutions (note the different scales on ordinate axes). **Left:** a 10ms pulse. **Right:** a 150 ms pulse. The convolution with a 150 ms pulse has peak responses that are saturated for each channel (i.e. convolution with a 200 ms pulse would not produce responses of greater magnitude). These peak responses were used to normalize N(t) for each channel.

The result of this convolution, **N(t)**, which represents the neural response to a given timecourse, is then multiplied by each pixel value of the spatially filtered images to produce a three dimensional array, **B(x,y,t)**. Each "slice" of this array represents the spatially filtered image at a particular moment in time, scaled by the value of the neural response, **N(t)**, at that time.

$$B_M(x,y,t)=F_M(x,y)\cdot N_M(t)$$ **Equation 4.5a**

$$B_P(x,y,t)=F_P(x,y)\cdot N_P(t)$$ **Equation 4.5b**

Next, M and P contrast gain functions are applied to each image, to obtain $\mathbf{D_M(x,y,t)}$ and $\mathbf{D_P(x,y,t)}$.

$$D_M(x,y,t)=\frac{66.5\cdot B(x,y,t)^{2.9}}{0.076^{2.9}+B(x,y,t)^{2.9}}$$ **Equation 4.6a**

$$D_P(x,y,t)=\frac{133.5\cdot B(x,y,t)^{0.88}}{2.8^{0.88}+B(x,y,t)^{0.88}}$$ **Equation 4.6b**

These equations are taken directly from Tapia & Breitmeyer (2011), who psychophysically measured the priming effects of masked and unmasked primes as a function of prime contrast. Note that these contrast images are no longer normalized to a peak value of 1.

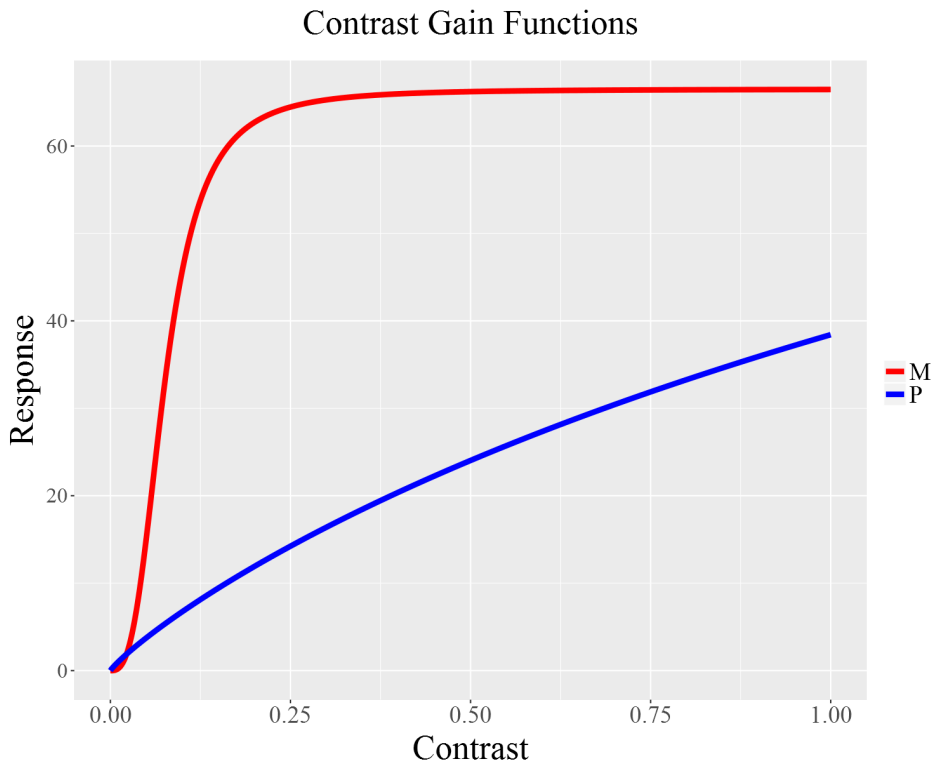

Figure 4.4. Contrast gain functions for the magno and parvocellular components of the model.

To simulate masking at a spatially invariant, object level, we included **$D_O(x,y,t)$**. This component represents a field of uniform contrast, temporally filtered by the P response of the mask.

$$N_O(t) = I_P(t) * T(t)_{Mask} \qquad \textbf{Equation 4.7}$$

$$D_O(x,y,t) = C(x,y)_{Object} \cdot N_O(t)_{Mask}, \text{ where } C(x,y)_{Object} = 1 \text{ for all x \& y} \qquad \textbf{Equation 4.8}$$

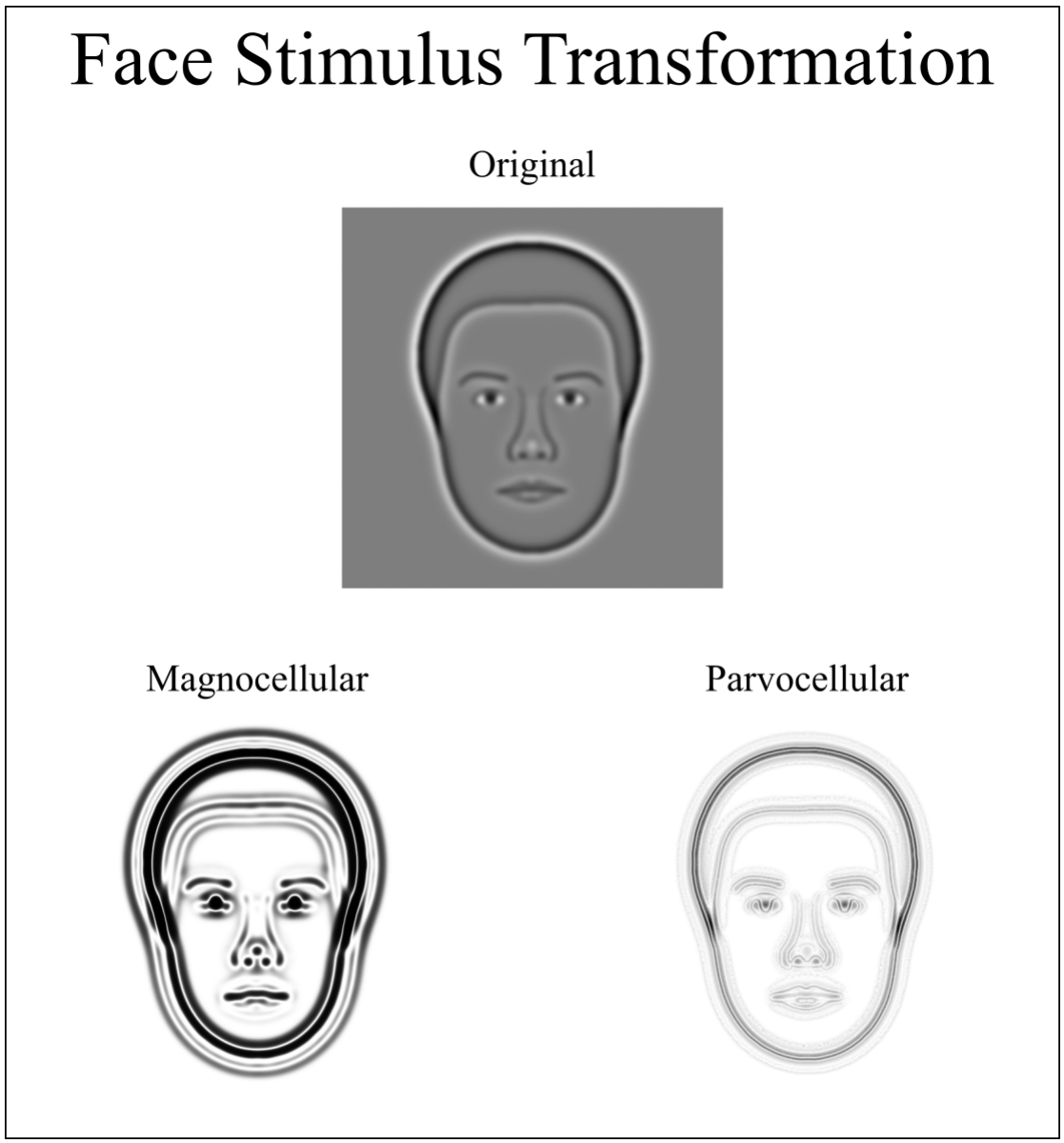

Figure 4.5. A visualization of the transformation from the original face stimulus into the magnocellular and parvocellular components, after undergoing the operations described in Equations 4.1, 4.2, & 4.6. For clarity, the bottom two images are rendered in reverse contrast and normalized.

Note that as $\mathbf{D_O(x,y,t)}$ is not spatially filtered, and does not depend on the size or shape of mask, this component is purely a function of $\mathbf{N_O(t)}$.

The final phase of the model combines the temporally and spatially filtered representations, $\mathbf{D(x,y,t)}$, of the target and mask to produce a visual response function, $\mathbf{VRF(t)}$, which is then averaged over time to produce a final response, $\mathbf{R}$. First, each slice of the P component of the target is point-wise divided by the point-wise sum of the mask's M and O components for the corresponding slice, to produce $\mathbf{V(x,y,t)}$. This operation can be understood as simulating M on P lateral inhibition, where the spatial spread of inhibition is defined by the receptive fields. The inclusion of the object component here simulates inhibition in a spatially homogeneous fashion, where all pixels of the target are inhibited equally by a uniform field whose strength over time varies according to Equation 4.7. Two scaling factors, $\mathbf{A_M}$ and $\mathbf{A_O}$, determine the relative strength that each of the two inhibitory components, $\mathbf{D_M(x,y,t)}$ and $\mathbf{D_O(x,y,t)}$ contribute.

$$V(x,y,t)=\frac{D_P(x,y,t)_{Target}}{1+A_M \cdot D_M(x,y,t)_{Mask}+A_O \cdot D_O(x,y,t)} \qquad \textbf{Equation 4.9}$$

Next, the spatial mean of each slice in $\mathbf{V(x,y,t)}$ is taken, to give $\mathbf{VRF(t)}$. In this operation, only pixels with non zero contrast values were included. This was done to ensure that the visual response function would not be more or less "diluted"

depending on the ratio of foreground vs. background pixels in the target image. Finally, **VRF(t)** is averaged over 300 ms to produce **R**.

$$R = \frac{\sum_{t=1}^{300} VRF(t)}{300}$$ **Equation 4.10**

4.2 Model Performance

Results from the model are shown below. Masking is plotted as the reciprocal of the model output ($1/\mathbf{R}$).

The two scaling parameters were set as follows:

$\mathbf{A_M} = 5 \times 10^5$, $\mathbf{A_O} = 0.8$.

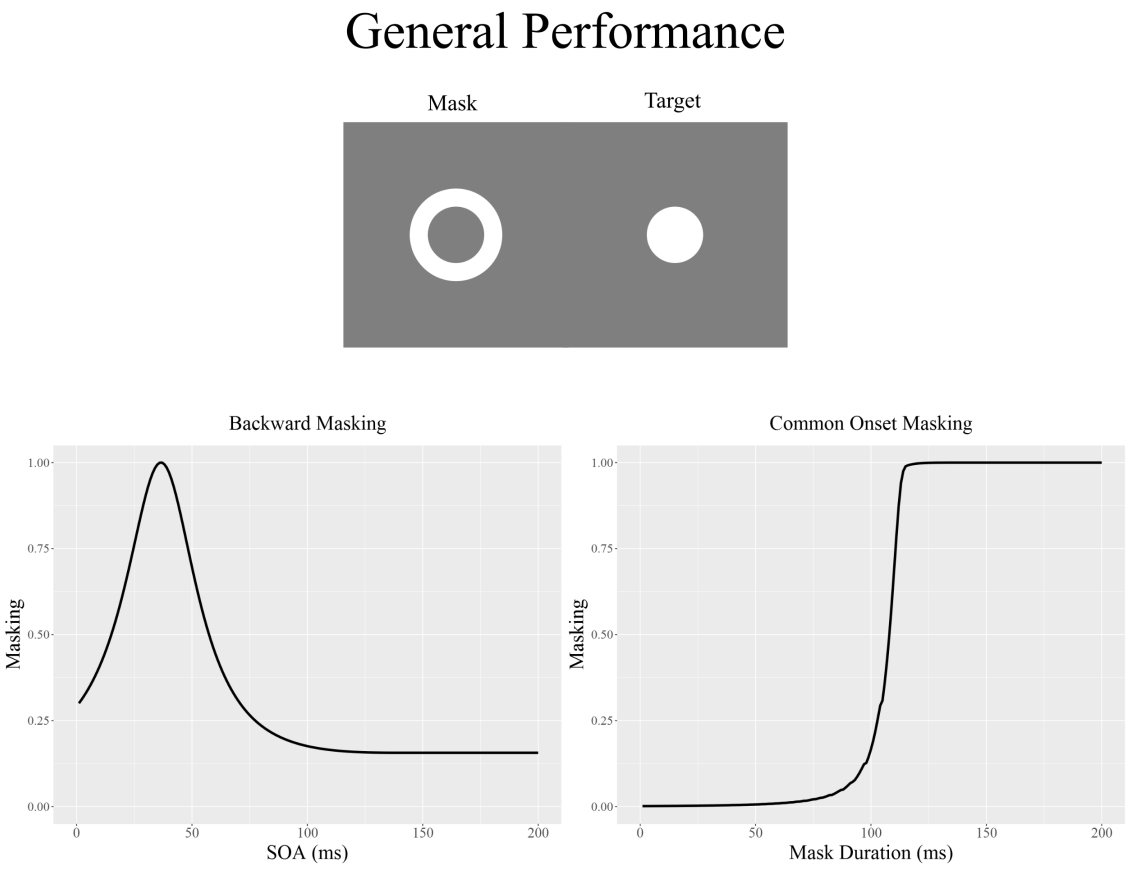

Figure 4.6. Backward and common onset masking functions using a disc and annulus, shown at **top**. For the backward masking simulations, the target and mask were each presented for 10 ms. In the common onset masking simulations, the target was presented for 10 ms. Note that masking is normalized in each plot.

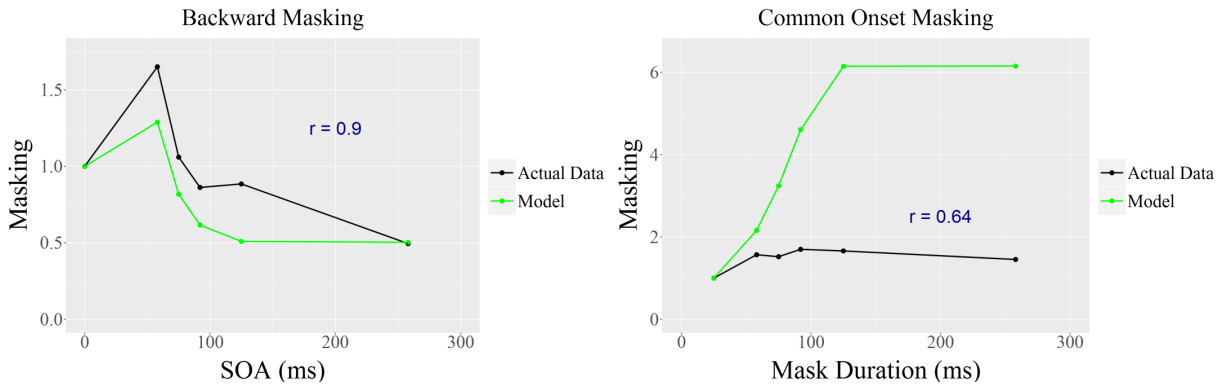

Figure 4.7. Model output vs. face masking data from Experiment 2A (Chapter 2). Data are plotted relative to the SOA 0 ms condition. Target and mask each had a duration of 25 ms. Note the different ordinate scales across the plots. Pearson correlation coefficients are shown for each of the two types of masking.

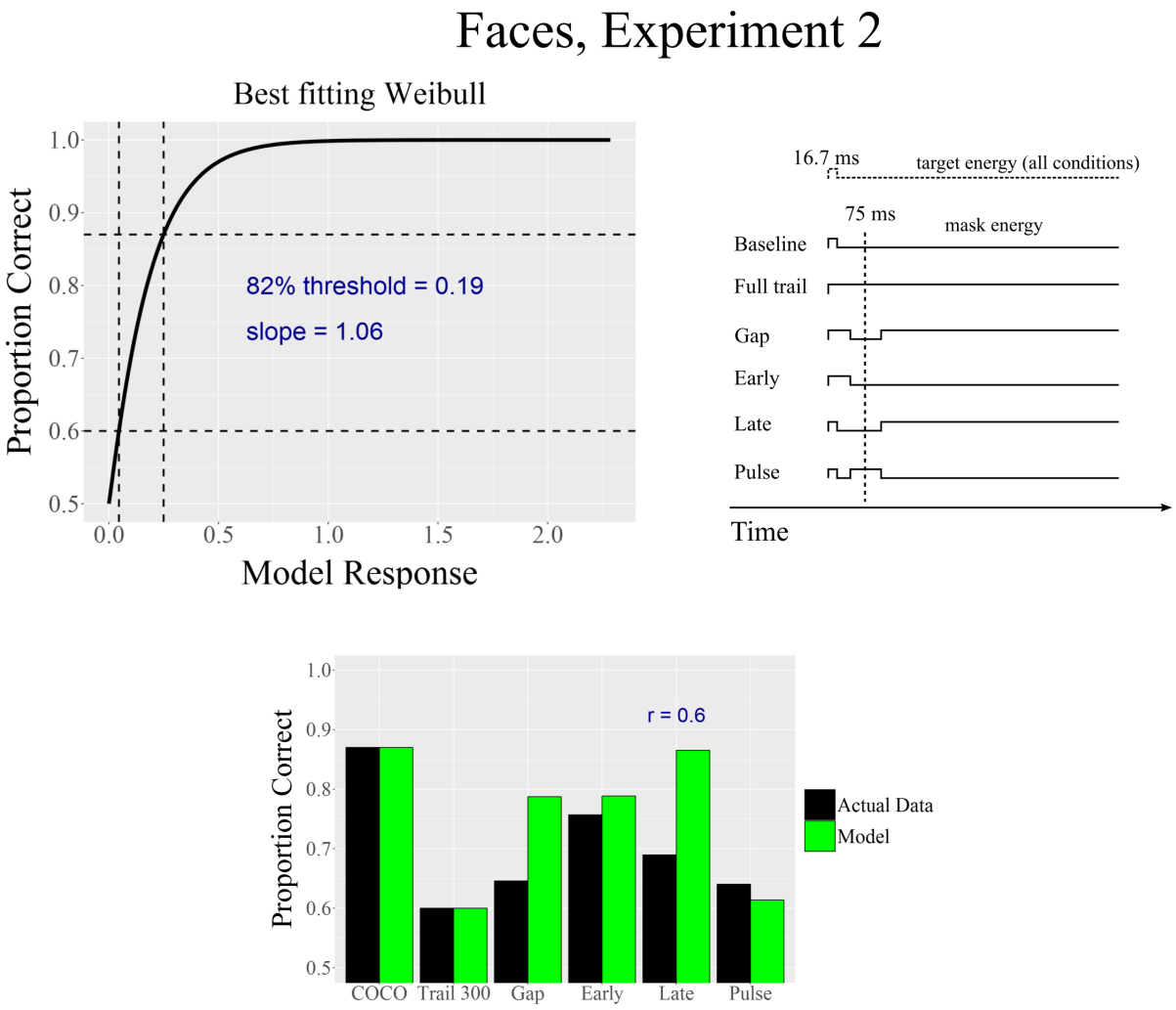

Figure 4.8. Model output vs. face masking data from Experiments 2C & 2D (Chapter 2). **Top Right:** Schematic timecourses shown for the various masking conditions. **Bottom:** Actual data vs. model data. The correlation here includes all conditions except for the Full Trail condition. Values of R in the model were converted to proportion correct values by first fitting a Weibull to the COCO (common onset common offset baseline) and Full Trail data. This Weibull was then used to generate the proportion correct values for the rest of the conditions. The fitted Weibull is shown in the **Top Left**. The horizontal dashed lines indicate the actual proportion correct values for the Baseline and Full Trail data, and the vertical dashed lines indicate the model's response to these two simulated conditions.

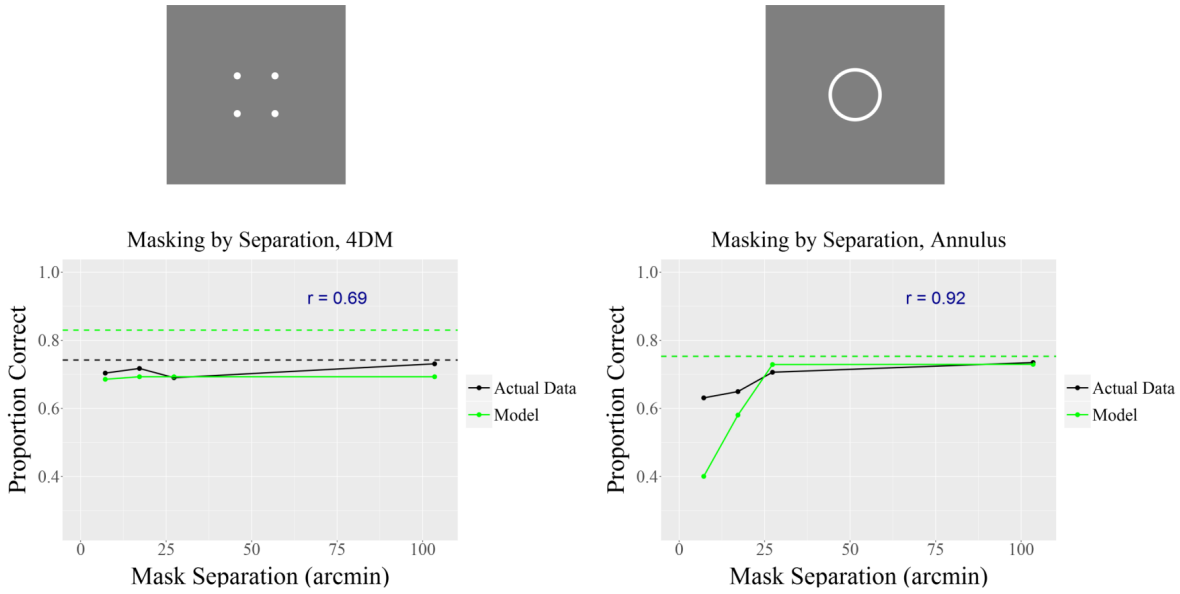

Figure 4.9. Model output vs. Experiment 3B (Chapter 3). To generate proportion correct values, a Weibull was first generated based on data from two conditions in Experiment 3A. These conditions were the 250 ms four-dot mask (4DM), and the common onset common offset four-dot mask (baseline). This Weibull was then used to transform model responses in the 4DM separation conditions. As the annulus mask is different from the 4DM, a new Weibull was generated for the annulus conditions. This was done by retaining the slope parameter of the 4DM Weibull, and finding a threshold parameter that, together with the slope parameter, fit the baseline data from the separation experiment (dashed lines). This is why, in the annulus condition, the baselines for both plots (actual, model) are essentially identical. Correlations shown here include all five conditions from each mask type (baseline + four separations).

4.3 Discussion

The model and the simulations described above yielded a respectable fit to data across the experiments described in Chapters 2 and 3. While there are shortcomings in some of the fits, it is noteworthy that with only two free parameters, the mean correlation across the experiments is 0.75. The classic type B backward masking function (Figure 4.6, left) is explained by the latency difference between the peak responses of the M and P channels (Figure 4.2). As described in Chapter 1.3, this latency difference means that the M component of the mask maximally inhibits the P component of the target when the target is presented before the mask, yielding the type B function seen here. The common onset masking function (Figure 4.6, right), where masking saturates after a sufficiently long mask duration, resembles a typical common onset masking function. Here, the saturated masking can be explained by two properties of the model. First, the M and P components of a stimulus reach peak saturation after a certain duration (Figure 4.3, right), which means that the peak M response of a mask will not grow arbitrarily large with increasing mask durations. Second, the P response of a brief target eventually decays (e.g. with a 10 ms pulse, the P response starts to decay at around 100 ms; see Figure 4.3, left). Because of the way masking is modeled here (Equation 9), the influence of an inhibitory M component of the mask at any particular point in time depends upon the strength of the P component of the target at that point in time. Thus, after the P component of the target has decayed completely, the presence of a mask cannot drive further masking, as there is nothing left to inhibit.

There are two notable instances where the model deviated from the empirical data, and both of these arise in the face masking experiments. The first is the relative magnitudes of peak masking in backward and common onset masking (Figure 4.7), and

the second is the relationship between the **Early** and **Late** conditions (Figure 4.8). While our experimental data showed roughly equal peak masking across backward and common onset masking conditions, the model yielded peak common onset masking that was over four and a half times greater than peak backward masking. This discrepancy may be related to the methodology of the experiment, where masking was quantified as face geometry thresholds in a two alternative forced choice task. The model presented here quantifies masking as proportional to the strength of the neural response to the target, which is probably more closely related to a criterion content such as brightness. Bischof & Di Lollo (1995) found that peak common onset masking was of a much greater magnitude than peak backward masking. In their study, observers had to report which of four sides of a square contained a gap, so it is unclear whether this task depended upon figural aspects of the stimulus, or simply upon brightness detection, but their data demonstrate that peak masking can vary depending upon whether masking is generated through backward vs. common onset masking. It is also worth noting that with some combinations of the two scaling factors, $\mathbf{A_M}$ and $\mathbf{A_O}$, it was possible to obtain peak backward and common onset masking values that were much closer together, however this had a negative impact upon other fits. The second discrepancy can be seen by examining Figure 4.8. The experimental data shows that the **Late** condition produced more masking than the **Early** condition. However, the model simulation showed the opposite, where **Early** showed higher masking than **Late**. This is because in the model, the latter pulse of the **Late** condition (which starts at about 100 ms) does not contribute as much to masking as the pulse in the **Early** condition, and this in turn is because the target's P component has already started to decay around 100 ms, and is therefore less vulnerable to masking, as discussed above. It may be the

case that the visual system may not weigh these time periods in the same way the model does. It is possible that if a secondary parvocellular reverberation of the target was included in the model, perhaps reflecting reentrant activity, then the **Late** condition would show masking more in line with the data, although it is unclear what effect this would have on the relative magnitudes of peak backward and common onset masking.

During the development of the model, versions were tested that included a (spatially filtered) P on P inhibitory component, as this is a key feature of the sustained transient dual channel model (Chapter 1.3). However, this inclusion did not appreciably improve the model fits, so I chose to omit this component in the interest of model simplicity. The object level inhibitory component is a form of P on P inhibition, however, and its inclusion was necessary to produce acceptable fits, particularly in the four-dot masking experiment.

In addition to the simulations described above, the model was tested against two other sets of data. The first is from Breitmeyer (1978), where the effects of mask duration upon backward masking were examined. In that study, increasing mask duration had the effect of shifting the masking function from type B towards type A (see Chapter 1.3 for a discussion of this finding). In my model, the spatial and temporal parameters of the stimuli were reproduced, as was the contrast. The two scaling factors were unchanged from those used in the above simulations. The results are shown below, in Figure 4.10. The most salient discrepancy is that in the model simulation, peak masking grows more rapidly as a function of mask duration than in Breitmeyer's data, a property that is related to the discrepancy between peak backward and common onset masking discussed earlier. However, two important features are shared. Peak masking shifts to smaller SOAs as mask duration increases, and peak masking increases in

magnitude. At even higher mask durations than the ones tested here, the model does indeed yield type A functions. The second set of data which was tested is described in Chapter 5.3 (see Figure 5.1). There, the model provides a very good fit to backward masking data from Francis & Cho (2007).

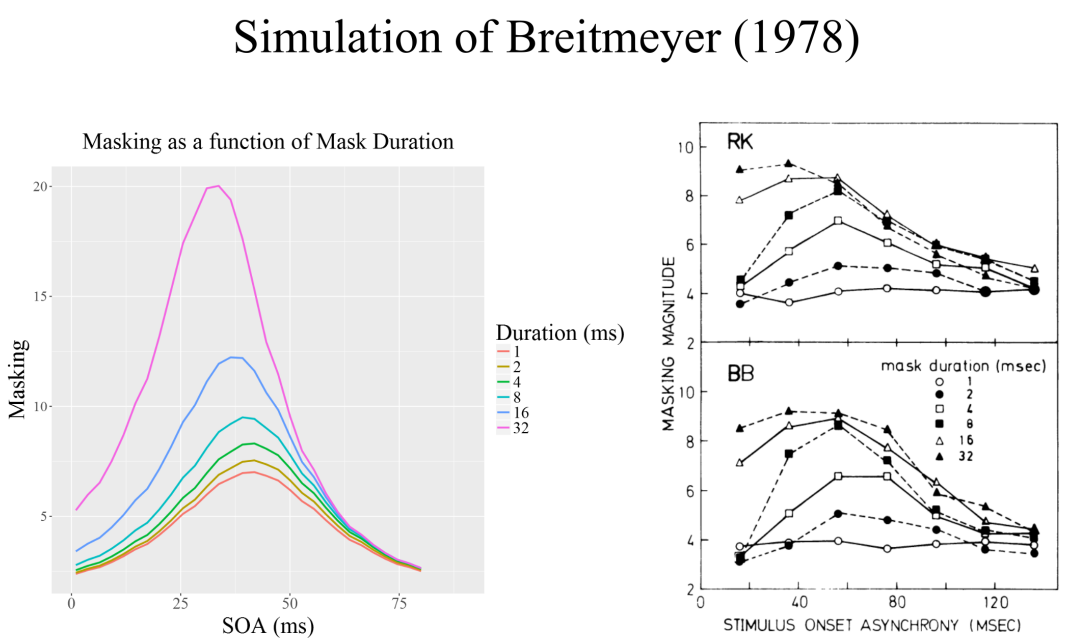

Figure 4.10. Left: Backward masking simulations as a function of mask duration. Stimuli are described in Breitmeyer (1978) and are the same ones as shown in Figure 4.6. The stimulus dimensions, contrast, and timings are replicated from the original paper. These results provide a qualitative match to the data found in Breitmeyer (1978). See Figure 1 from their paper, shown here on **right**, which shows data from two observers. Note that in both their data and our model, as mask duration increases, peak SOA becomes smaller, and peak masking increases.

The computational model described here is flexible enough to test a variety of masking conditions with explicitly defined spatial and temporal stimulus properties. It performs reasonably well overall on the tested data sets, and captures many salient

properties of the data. Perhaps the most important finding is that the inclusion of an object level inhibitory component was critical in providing good fits for the target mask separation data (Figure 4.9). This lends support to the idea that masking is not a unitary process, and that a subset of the involved mechanisms shows a high degree of spatial invariance.

Chapter Five

General Discussion

5.1 Summary of findings

This dissertation describes a series of experiments designed to measure masking under a set of parameters that are either not typically tested together (e.g. common onset masking in central visual field), or are not directly compared (e.g. different mask types under a common onset presentation; different types of schedules using the same spatial stimulus properties). In the first set of experiments, a single class of stimuli (faces) was measured under a variety of temporal conditions, including backward and common onset masking. Here, we found that a common onset, delayed offset (trailing) mask disrupted target identification to the same degree as a briefly flashed backward mask. Based on this initial result, one might expect that the key property shared by both conditions was the physical presence of the mask at a certain moment (i.e. 58 ms after target onset). However, in a third condition, we introduced a temporal gap which was centered around this moment, in a common onset trailing mask. Here, masking was not reduced at all. By testing a variety of other temporal schedules, we were able to account for our data using a simple computational model that incorporated two key ideas. The first is that the mask leaves a trace in the visual system that decays after it physically offsets (i.e. visual persistence). The second is that there is a distinct process, time locked to target onset, that interacts with this trace to produce masking. While the limitations of this model are significant, and the findings are not surprising, it is precisely this use of a variety of temporal schedules that allowed for such a model to be

developed. Specifically, by testing the visual system's response to a set of conditions that went beyond simply varying SOA, we had more information with which to create a model. To our knowledge, there are only two precedents to this approach (Di Lollo et al., 2004; Francis & Hermens, 2002), who tested existing models under both backward and common onset masking conditions, and while some of the simulations in these studies are more sophisticated than the one used here, they only compared two general forms of temporal schedule: varying the SOA, and varying the duration of a common onset trailing mask (in the case of Di Lollo et al., the mask onset immediately after target offset). In our experiments, we compared a number of other conditions, for example conditions with multiple mask flashes within a single trial.

Our next set of experiments revealed a number of important findings relating to OSM. Here, we first replicated Filmer et al.'s (2015) finding that OSM can be reliably obtained in central visual field. We then directly compared a four-dot mask with an annulus mask under common onset conditions, and examined how masking with each mask type varied as a function of target mask separation. In the last experiment, we tested how a four-dot mask affected orientation judgments using an orientation matching task. The data from these experiments lend good evidence that OSM in central visual field reflects mechanisms that have a high degree of spatial invariance. While Filmer et al.'s demonstration of masking in central visual field using a sparse four-dot mask is compatible with a spatially invariant mechanism, the current experiments represent the first direct test of a sensitivity to spatially local interactions, and provide novel evidence in favour of a spatially global mechanism. In particular, the finding that target mask separations of up to at least half a degree of visual angle does not attenuate masking when a four-dot mask is used closely matches the data found by (Di Lollo et al., 2000), and supports the idea that the masking found in ours and Filmer

et al.'s work is indeed akin to the more traditional, parafoveal demonstrations of OSM. It also directly addresses an objection raised in response to Di Lollo et al. This objection (Breitmeyer & Öğmen, 2006) states that the observed insensitivity to contour proximity (between target and mask) may be due to the increased receptive field sizes found with increased retinal eccentricity (Dacey & Petersen, 1992; Freeman & Simoncelli, 2011), and that the masking that Di Lollo et al. found was not necessarily due to object level mechanisms. The experiments reported here support Di Lollo et al.'s account on two fronts. First, our stimuli were centrally presented. Our target mask separation of 27.3 arcmin (which showed no attenuation of masking relative to a separation of 7.2 arcmin) represents a mask eccentricity of just over a degree of visual angle. In comparison, Di Lollo et al.'s data come from mask eccentricities as high as almost four degrees. Second, we found strong attenuation of annulus masking as a function of target mask separation. If large receptive field sizes were obscuring the effect of any contour interactions at the presented eccentricities, then such an effect should have been similarly obscured with the annulus mask, yet the attenuation effect was striking with the annulus mask. This second piece of evidence demonstrates the value of directly comparing the effect of parameters which, due to methodological clustering (Chapter 1.6), are not typically studied together (in this case, comparing a four-dot and an annulus mask in a common onset paradigm).

Another possibility worth exploring is that while our data reflect object based mechanisms, they may not necessarily reflect spatially invariant mechanisms. It is known that attentional inhibitory mechanisms can spread across an object (Jordan & Tipper, 1999), and that the centroid of an object can serve as the locus of attention (Zhou, Chu, Li, & Zhan, 2006). Given these two facts, it may be the case that the four dot mask, as an object, inhibited the target through spatially sensitive interactions

(between the centroids of the mask and target), which then spread across the entire target, including the oriented bar. If this were true, then one might expect no substantial change in masking as the four dots moved further away from the target, since the *centroid* of the four dots remained in the same location. It would also account for the finding that cardinal and oblique orientations were equally vulnerable to masking, despite the different separations of these sets of orientations from the mask. The importance of a centroid is is also consistent with Jiang & Chun's (2001) finding that four dot masking decreased with distance from the target, when the dots were moved *as a whole*—in their study, the centroid moved away from the target with increasing target mask separations. It should be noted, however, that the smallest separation increment they used was about 1 degree. It is unclear whether they would have found an attenuation of masking had they shifted the four dot mask away from the target by half a degree. It's also worth noting that they found masking with the highest separation tested (3.75 degrees), suggesting the involvement of larger receptive field sizes found in higher visual areas. The strong attenuation of masking we found with the annulus mask may be due to the relatively high degree of classical surround suppression that would likely be expected with the increased opportunity for local interactions found between the parallel contours of the mask and target annuli (with spreading inhibition disrupting the bar component of the target). Additionally, the featural similarity between the annuli may have boosted this suppression (e.g. see Flevaris & Murray, 2015; Habak, Wilkinson, & Wilson, 2006). Follow up experiments, for example using a single bar as a target, rather than a bar embedded in an annulus, may resolve some of these questions, and may help shed light on whether these interactions are occurring in early stages, such as V1, or in higher stages, such as V4 and beyond.

Our target mask separation data also shed light on a question about the first set of face masking experiments. In those earlier experiments, it was unclear how much of the masking effect was due to object level vs. lower level mechanisms. In Experiment 2A, masking increased as a function of mask duration, up until around 100 ms (Figure 2.5). While this may have been due to OSM mechanisms involving reentrant processing, it may also have been due to an increased amount of intrachannel sustained on sustained inhibition associated with the increased mask energy that occurs with longer mask durations (Chapter 1.3). The strong attenuation of annulus masking with increased target mask separations suggests that with the annulus mask, spatially local mechanisms played a dominant role in producing masking. Given the nature of the face stimuli (e.g. the head outline is somewhat similar to an annulus), it is possible that such mechanisms similarly played a significant role in our face masking experiments, although see Loffler et al. (2005). Another consideration here is that the temporal dynamics of visual processing may differ between face stimuli and more simple shapes, such as an annulus. Electrophysiological work by Sugase, Yamane, Ueno, & Kawano (1999) found that individual face selective neurons in macaque cortex encoded information about global features (e.g. whether the stimulus is a face or not, or what species the face belongs to), followed by information about fine features (facial identity, expression). Importantly, there was a delay of about 51 ms between these two stages of processing, a result consistent with reentrant accounts of visual processing. This was followed up by work investigating the response of populations of neurons, and again, a latency difference of about 50 ms was found between the encoding of global and fine information (Matsumoto et al., 2005). In light of these studies, it would be interesting to compare common onset masking between a simple annulus and a face, as a function

of mask duration, as the former stimulus contains only global information, while the latter contains both global and fine information.

5.2 Limitations of computational model

In Chapter 4, I presented a computational model that was developed to account for findings from both sets of experiments. Despite only having two free parameters (the scaling factors $\mathbf{A}_{\mathrm{M}}$ and $\mathbf{A}_{\mathrm{O}}$), the model provided a respectable fit to the data. However, there are a few important limitations. First, the model does not account for integrative effects such as contrast reduction due to luminance summation, camouflage, or degradation that occur when the target and mask occur close together in time (Chapter 1.2). Rather, the only mechanisms of masking that are modeled are lateral and global inhibition of the target by the magnocellular and parvocellular mask components, respectively. Second, while the latter component represents masking at an object level (i.e. it is spatially invariant, or global), it is assumed to occur independent of any similarity between target and mask. In this sense, our model more closely resembles object substitution than object updating. Third, performance is modeled by a single linking assumption between the visual response function (VRF) and performance (namely, the integral of VRF over time), and the VRF is explicitly tied to the parvocellular response of the target. While this may be a reasonable approach to modeling a criterion content such as brightness, it is not clear that a single criterion was used by observers across the set of experiments described in this dissertation, even if we restrict this to the set of experiments that were included in the model. As we have seen, however, task parameters (which determine which criteria observers use) can have significant impact on masking functions (Chapter 1.3), and the use of a single linking assumption limits the generalizability of our model. Finally, our model does not take

into account the noise mask used in Chapter 3, nor does it predict that a noise degraded target is required for masking. Despite the model's limitations, it performed reasonably well when compared against the data across both sets of experiments, which represented a diverse set of temporal and spatial conditions. Furthermore when tested with the conditions used in Breitmeyer (1978), the model output provided a qualitative match (Figure 4.10).

5.3 Does the model implement mask blocking?

Francis & Cho (2007) describe a computational principle called mask blocking that can yield type B shaped backward masking functions (see also Francis, 2000). This principle operates under a framework involving a set of differential equations describing three variables: the target signal, the mask signal, and a visual response function (VRF). The derivative of VRF (with respect to time) is proportional to the current value of VRF, and to the strength of the target and mask signals, which contribute towards growth and decay of VRF, respectively. At time values when the mask signal is weaker than the target signal, the net effect on the derivative of VRF is one that encourages growth. It follows that if the stimulus intensity of the mask is relatively weak compared to that of the target, the effect of the mask, upon the VRF, will be blocked when the target and mask appear together. Once the target signal has decayed sufficiently, however, presentation of the mask will produce a mask signal that is greater than that of the target signal. This dynamic yields a type B function, where peak masking occurs when the mask is presented after target offset. Francis showed how a number of masking models operate through this mask blocking principle. Furthermore, Francis and Hermens (2002) showed how these models, which can produce type B backward masking functions, can also be modified to provide a good match to the

common onset masking data found in Di Lollo et al. (2000), where masking increased with mask duration, and interacted with set size. However, these models did not fare well when tested against experimental data from Francis & Cho, who tested observers in two backward masking experiments. In the first, an annulus mask was used, and in the second, a sparse four-dot mask was used. The mask blocking principle predicts that the four-dot mask would produce a type B masking function, as here, the mask is relatively weak (Francis, 2000; Francis & Herzog, 2004); and indeed, Francis & Cho's simulations of the three models that use this principle (Bridgeman, 1978; Francis, 1997; Weisstein, 1972) showed type B functions with a four-dot mask (as the models do not allow explicit representations of the spatial extent of stimuli, this type of mask was emulated by setting mask intensity to a tenth of that of the target). The empirical data, however, showed a type B function with the annulus mask, and a type A function with the four-dot mask, demonstrating that it is unlikely that mask blocking gives a sufficient account of the observed masking phenomena. To see how my own model (Chapter 4) fared against the empirical data in Francis & Cho's study, I simulated the conditions of the two experiments. As Figure 5.1 shows, the model output matches the main empirical finding from Francis and Cho: as the mask becomes sparser, backward masking shifted from a type B to a type A function. In my model, this behaviour can be explained by considering the two inhibitory components: the spatially local inhibition of the mask's magnocellular component upon the target's parvocellular component, and the spatially global inhibition of the mask's parvocellular component upon the target's parvocellular component. As the M component peaks earlier than the P component (Figures 4.2, 4.3), the spatially local M on P inhibitory component will peak when the target appears before the mask. The P on P inhibition, however, tends to get stronger when mask and target appear closer in time (if the mask and target duration are

identical, then this P on P inhibition will peak when mask and target onset together). Importantly, as the P on P inhibition in my model is spatially invariant, the magnitude of this inhibition remains constant across the properties of the mask. A high contrast, spatially dense mask that is adjacent to the target will provide the same P on P inhibition as a low contrast, sparse mask that is far away from the target. This explains the pattern of results seen in Figure 5.1. When the mask is strong, the M on P component dominates, giving rise to type B effects. When the mask is weak, the P on P component dominates, giving rise to type A effects. It is clear that my model does not instantiate the principle of mask blocking, as mask blocking predicts the opposite outcome. It is also worth noting that in Francis & Cho's simulations, the linking hypothesis between VRF and perceptual awareness was very similar to the one used in my own model. That is, perceptual awareness (quantified as **R** in my model) is proportional to the integral, over time, of VRF. This need not be the case, however. For example, the percept strength could be modeled as the *duration* of the VRF when it is above a certain threshold, and modeling in this fashion can produce subtle differences in masking properties (c.f. theorems 1 & 2 in Francis, 2000). Thus, differences between the Francis & Cho's simulations and my own cannot be explained by different linking hypotheses, and are instead more likely due to different underlying computational principles.

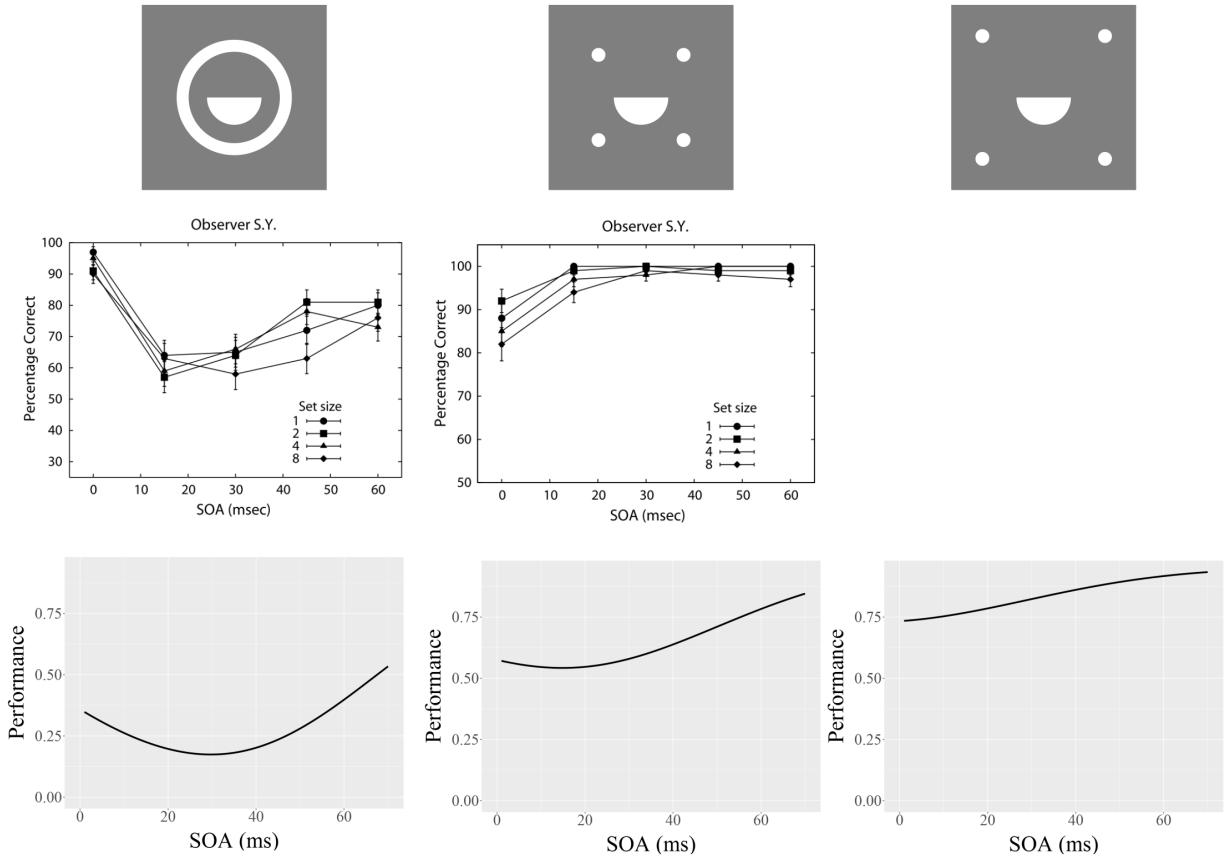

Figure 5.1. Comparison of output of computational model (Chapter 4) with data from Francis & Cho (2007). At **top** are shown three different target and mask configurations, the **left** and **center** of which were used in Francis & Cho's study. In the **middle row** are shown backward masking functions for a single observer from Francis & Cho (2007) associated with these stimuli (reproduced from Francis & Cho, 2007). At **bottom** are shown masking functions generated by my computational model with stimuli whose dimensions, contrast, and timings were replicated from Francis & Cho's study. One difference was that in their study, the stimuli were part of multistimulus array; however, the set size 1 conditions are comparable to my own simulations. In both their study and my own simulations, the same pattern is seen: as the mask becomes weaker, masking shifts from type B to type A. As my simulation didn't produce a "pure" type A function with the four-dot mask used in their study (see bottom middle plot), I tested my model with an even weaker mask, by increasing target mask separation (see top right). Here, a monotonic type A backward masking function was obtained. The "performance" plotted in my simulations reflect raw model outputs (R).

5.4 Future directions and final thoughts

There are a few immediate questions that arise out of this dissertation. Why is an initial forward mask needed to facilitate OSM in a foveated and attended target, and does it need to be forward masked, or would a noise mask presented simultaneously with the target achieve the same effect? In Chapter 3.5, I presented some ideas about how a noise mask might introduce a requirement for perceptual filling in processes, which in turn may ripen the conditions for OSM to occur. These ideas could potentially be tested by trying to recreate OSM with different targets and forward masks. For example, a small OSM effect in central visual field was found in Dux et al. (2010) when a stream of digits was presented prior to the target and four-dot mask, even in a condition where observers were told to ignore the digits and not do any arithmetic on them. While this may have been due to an inadvertent increase in cognitive load, it may also have been due to the target being forward masked by the most recent digit. If this is the case, then any shape with sufficient energy that is presented at this location should produce the same effect. A recent study (Filmer, Wells-Peris, & Dux, 2017) provides some answers here. In Experiment 1 of their paper, they obtained reliable OSM in central visual field when the target consisted of a noise degraded shape, and without the use of a forward mask. Furthermore, comparing the results from their second and third experiments suggests that the presence of a digit, 100 ms before the target, slightly increased OSM due to forward masking. Thus, while forward masking may play a minor role in producing OSM in central visual field, the key condition seems to be that the target be degraded by noise. A related question is whether the target representation in OSM is completely abolished, or whether it is only partially degraded or reduced in apparent contrast. My own results suggest the former; however, as discussed in Chapters 3.4.3 & 3.5, it is possible the latter mechanisms could be

responsible for degrading or reducing contrast below a threshold necessary for conscious perception. While the perceptual outcomes of these two possibilities are the same, the means by which this occurs is an important distinction, and deserves further consideration, especially in light of the contrasting findings of Harrison et al. (2016). Shedding light on these questions is important if we are to improve our understanding of how the phenomenon of OSM relates to the object updating framework (Goodhew, 2017).

This dissertation, in particular the computational model presented in Chapter 4, supports the growing idea that masking can involve multiple processes (Albrecht & Mattler, 2016; Breitmeyer, 2015; Reeves, 1982; Turvey, 1973). A more sophisticated model that attempts to model a larger number of these processes in greater detail may prove useful. Such a model, while highly ambitious, would presumably be able to successfully predict masking effects across a variety of temporal schedules, stimuli, and tasks. This dissertation demonstrates that combining parameters that are typically restricted to different paradigms can bear fruit. It also generates questions for further exploration that can deepen our understanding of masking and the visual system.

References

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

Albrecht, T., & Mattler, U. (2016). Individually different weighting of multiple processes underlies effects of metacontrast masking. *Consciousness and Cognition*, *42*, 162–180.

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), 646–61.

Averbach, E., & Sperling, G. (1961). Short term storage and information in vision. In C. Cherry (Ed.), *Information theory* (pp. 196-211). Washington, DC: Butterworth & Co.

Bacon-Macé, N., Macé, M. J.-M., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: backward masking and natural scene categorisation. *Vision Research*, *45*(11), 1459–69.

Baker, D. H., Kaestner, M., & Gouws, A. D. (2016). Measurement of crosstalk in stereoscopic display systems used for vision research. *Journal of Vision*, *16*(15), 14-14.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Schmidt, A. M., ... Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–54.

Benardete, E. A., & Kaplan, E. (1997). The receptive field of the primate P retinal ganglion cell, I: Linear dynamics. *Visual Neuroscience*, *14*(1), 169–85.

Bischof, W. F., & Di Lollo, V. (1995). Motion and metacontrast with simultaneous onset of stimuli. J*ournal of the Optical Society of America. A, Optics, Image Science, and Vision*, *12*(8), 1623–36.

Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of

stimuli including White, Howe and simultaneous brightness contrast. *Vision Research*, *45*(5), 607–15.

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

Breitmeyer, B. G. (1978). Metacontrast masking as a function of mask energy. *Bulletin of the Psychonomic Society*, *12*(1), 50–52.

Breitmeyer, B. G. (1992). Parallel processing in human vision: History, review, and critique. *Advances in psychology*, *86*, 37-78.

Breitmeyer, B. G. (2007). Visual masking: past accomplishments, present status, future developments. *Advances in Cognitive Psychology*, *3*(1–2), 9–20.

Breitmeyer, B. G. (2014). Contributions of magno- and parvocellular channels to conscious and non-conscious vision. *Philosophical Transactions of the Royal Society of London*, *369*(1641), 20130213.

Breitmeyer, B. G. (2015). Psychophysical "blinding" methods reveal a functional hierarchy of unconscious visual processing. *Consciousness and Cognition*, *35*, 234–250.

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., Kafaligönül, H., Öğmen, H., Mardon, L., Todd, S., & Ziegler, R. (2006). Meta- and paracontrast reveal differences between contour- and brightness-processing mechanisms. *Vision Research*, *46*(17), 2645–2658.

Breitmeyer, B. G., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision.* Oxford: Oxford University Press.

Breitmeyer, B. G., Rudd, M., & Dunn, K. (1981). Metacontrast investigations of sustained-transient channel inhibitory interactions. *Journal of Experimental Psychology. Human Perception and Performance*, *7*(4), 770–779.

Breitmeyer, B. G., & Tapia, E. (2011). Roles of contour and surface processing in microgenesis of object perception and visual consciousness. *Advances in Cognitive Psychology*, *7*(1), 68–81.

Breitmeyer, B. G., & Williams, M. C. (1990). Effects of isoluminant-background color on metacontrast and stroboscopic motion: interactions between sustained (P) and transient (M) channels. *Vision Research*, *30*(7), 1069–1075.

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

Bridgeman, B. (2007). Common-onset masking simulated with a distributed-code model. *Advances in Cognitive Psychology*, *3*(1-2), 33–40.

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*(3), 121–138.

Burr, D. C. (1984). Summation of target and mask metacontrast stimuli. *Perception*, *13*(2), 183–92.

Camp, A. J., Cheong, S. K., Tailby, C., & Solomon, S. G. (2011). The impact of brief exposure to high contrast on the contrast response of neurons in primate lateral geniculate nucleus. *Journal of Neurophysiology*, *106*(3), 1310-1321.

Camp, S. J., Pilling, M., Argyropoulos, I., & Gellatly, A. (2015). The role of distractors in object substitution masking. Journal of Experimental Psychology. *Human Perception and Performance*, *41*(4), 940–57.

Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., & Gross, J. (2011). Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS ONE*, *6*(1), e14543.

Chakravarthi, R., & Cavanagh, P. (2009). Recovery of a crowded object by masking the flankers: determining the locus of feature integration. *Journal of Vision*, *9*(10), 4.1-9.

Choo, H., & Franconeri, S. L. (2010). Objects with reduced visibility still contribute to size averaging. *Attention, Perception & Psychophysics*, *72*(1), 86–99.

Connors, B. W., Malenka, R. C., & Silva, L. R. (1988). Two inhibitory postsynaptic potentials, and GABAA and GABAB receptor-mediated responses in neocortex of rat and cat. *The Journal of Physiology*, *406*, 443–68.

Crawford, B. H. (1947). Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society of London B: Biological Sciences*, *134*(875), 283-302.

Daar M., Wilson, H. R. (2015). Masking with faces in central visual field under a variety of temporal schedules. *Vision Research*, *116*(2015), 1-12.

Daar, M., & Wilson, H. R. (2016). A closer look at four-dot masking of a foveated target. *PeerJ*, *4*, e2068.

Dacey, D. M., & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(20), 9666–9670.

De Valois, R. L., & Cottaris, N. P. (1998). Inputs to directionally selective simple cells in macaque striate cortex. *Proceedings of the National Academy of Sciences*, *95*(24), 14488-14493.

Di Lollo, V., Bischof, W. F., & Dixon, P. (1993). Stimulus-Onset Asynchrony Is Not Necessary for Motion Perception or Metacontrast Masking. *Psychological Science*, *4*(4), 260–263.

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.

Di Lollo, V., von Mühlenen, A., Enns, J. T., & Bridgeman, B. (2004). Decoupling stimulus duration from brightness in metacontrast masking: data and models. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(4), 733–45.

Dux, P. E., Visser, T. A., Goodhew, S. C., & Lipp, O. V. (2010). Delayed reentrant processing impairs visual awareness an object-substitution-masking study. *Psychological Science*, *21*(9), 1242-1247.

Edwards, V. T., Hogben, J. H., Clark, C. D., & Pratt, C. (1996). Effects of a red background on magnocellular functioning in average and specifically disabled readers. *Vision Research*, *36*(7), 1037–1045.

Elze, T. (2010). Misspecifications of Stimulus Presentation Durations in Experimental Psychology: A Systematic Review of the Psychophysics Literature. *PLoS ONE, 5*(9), e12792.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research*, *44*(12), 1321–31.

Enns, J. T., Lleras, A., & Moore, C. M. (2010). Object updating: a force for perceptual continuity and scene stability in human vision. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (pp. 503–520). Cambridge: Cambridge University Press.

Enns, J. T., & Lollo, V. Di. (1997). Object Substitution: A New Form of Masking in Unattended Visual Locations. *Psychological Science*, *8*(2), 135–139.

Eriksen, C. W. (1980). The use of a visual mask may seriously confound your experiment. *Perception & Psychophysics*, *28*(1), 89–92.

Eriksen, C. W., & Hoffman, M. (1963). Form recognition at brief durations as a function of adapting field and interval between stimulations. *Journal of Experimental Psychology*, *66*(5), 485–499.

Exner, S. (1868). Über die zu einer Gesichtswahrnehmung nöthige Zeit. *Wiener Sitzungsber Math-Naturwiss Cl Kaiserlichen Akad Wiss*, *58*(Part 2), 601-32.

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. a F. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience, 19*(9), 1488–97.

Fehrer, E., & Biederman, I. (1962). A comparison of reaction time and verbal report in the detection of masked stimuli. *Journal of Experimental Psychology*, *64*(2), 126–130.

Fehrer, E., & Raab, D. (1962). Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, *63*(2), 143–147.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47.

Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature*, *358*(6389), 756.

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*(6), 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.

Filmer, H. L., Wells-Peris, R., & Dux, P. E. (2017). The role of executive attention in object substitution masking. *Attention, Perception, & Psychophysics*, 1-8.

Flevaris, A. V., & Murray, S. O. (2015). Feature-based attention modulates surround suppression. *Journal of Vision*, *15*(1), 29-29.

Francis, G. (1997). Cortical dynamics of lateral inhibition: metacontrast masking. *Psychological Review*, *104*(3), 572.

Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review*, *107*(4), 768–785.

Francis, G., & Cho, Y. S. (2007). Testing models of object substitution with backward masking. *Perception & Psychophysics*, *69*(2), 263–75.

Francis, G., & Cho, Y. S. (2008). Effects of temporal integration on the shape of visual backward masking functions. Journal of Experimental Psychology. *Human Perception and Performance*, *34*(5), 1116–28.

Francis, G., & Hermens, F. (2002). Comment on "Competition for consciousness among visual events: the psychophysics of reentrant visual processes" (Di Lollo, Enns, & Rensink, 2000). *Journal of Experimental Psychology: General*, *131*(4), 594-596.

Francis, G., & Herzog, M. H. (2004). Testing quantitative models of backward masking. *Psychonomic Bulletin & Review*, *11*(1), 104–12.

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195–1201.

Ganz, L. (1966). Mechanism of the figural aftereffects. *Psychological Review*, *73*(2), 128–50.

Ghose, T., Hermens, F., & Herzog, M. H. (2012). How the global layout of the mask influences masking strength. *Journal of Vision*, *12*(13), 1-15.

Goodhew, S. C. (2017). What have we learned from two decades of object-substitution masking? Time to update: Object individuation prevails over substitution. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(6), 1249–1262.

Goodhew, S. C., Boal, H. L., & Edwards, M. (2014). A magnocellular contribution to conscious perception via temporal object segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 948-596.

Goodhew, S. C., Dux, P. E., Lipp, O. V, & Visser, T. a W. (2012). Understanding recovery from object substitution masking. *Cognition*, *122*(3), 405–15.

Goodhew, S. C., & Edwards, M. (2016). Object individuation is invariant to attentional diffusion: Changes in the size of the attended region do not interact with object-substitution masking. *Cognition*, *157*, 358–364.

Goodhew, S. C., Edwards, M., Boal, H. L., & Bell, J. (2015). Two Objects or One? Similarity Rather Than Complexity Determines Objecthood When Resolving Dynamic Input. Journal of Experimental Psychology. *Human Perception & Performance*, *41*(1), 102–110.

Goodhew, S. C., Greenwood, J. A., & Edwards, M. (2016). Categorical information influences conscious perception: An interaction between object-substitution masking and repetition blindness. *Attention Perception & Psychophysics*, *78*(4), 1186–1202.

Goodhew, S. C., Pratt, J., Dux, P. E., & Ferber, S. (2013). Substituting objects from consciousness: a review of object substitution masking. *Psychonomic Bulletin & Review*, *20*(5), 859–77.

Goodhew, S. C., Visser, T. a W., Lipp, O. V, & Dux, P. E. (2011). Implicit semantic perception in object substitution masking. *Cognition*, *118*(1), 130–4.

Green, M. (1981). Spatial frequency effects in masking by light. *Vision Research*, *21*(6), 861–6.

Growney, R., & Weisstein, N. (1972). Spatial Characteristics of Metacontrast. J*ournal of the Optical Society of America*, *62*(5), 690-696.

Growney, R., Weisstein, N., & Cox, S. I. (1977). Metacontrast as a function of spatial separation with narrow line targets and masks. *Vision Research*, *17*, 1205–1210.

Habak, C., Wilkinson, F., & Wilson, H. R. (2006). Dynamics of shape interaction in human vision. *Vision Research*, *46*(26), 4305–20.

Harrison, G. W., Rajsic, J., & Wilson, D. E. (2016). Object-substitution masking degrades the quality of conscious object representations. *Psychonomic Bulletin & Review*, *23*(1), 180-6.

Harrison, K., & Fox, R. (1966). Replication of reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, *71*(1), 162–3.

Heeley, D. W., & Timney, B. (1988). Meridional anisotropies of orientation discrimination for sine wave gratings. *Vision Research*, *28*(2), 337-344.

Chicago

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing and test-field luminances. *Journal of Experimental Psychology*, *50*(2), 89–96.

Ikeda, H., & Wright, M. J. (1975). Spatial and temporal properties of'sustained' and'transient' neurones in area 17 of the cat's visual cortex. *Experimental Brain Research*, *22*(4), 363–383.

Jannati, A., Spalek, T. M., & Di Lollo, V. (2013). A novel paradigm reveals the role of reentrant visual processes in object substitution masking. *Attention, Perception, & Psychophysics*, *75*(6), 1118-1127.

Jastrow, J. (1892). Studies from the University of Wisconsin: on the judgment of angles and positions of lines. *The American Journal of Psychology*, *5*(2), 214-248.

Jiang, Y., & Chun, M. M. (2001). The spatial gradient of visual masking by object substitution. *Vision Research*, *41*(24), 3121–31.

Jonas, P., & Buzsaki, G. (2007). Neural inhibition. *Scholarpedia*, *2*(9), 3286.

Jordan, H., & Tipper, S. P. (1999). Spread of inhibition across an object's surface. *British Journal of Psychology*, *90*(4), 495-507.

Kafaligonul, H., Breitmeyer, B. G., & Öğmen, H. (2015). Feedforward and feedback processes in vision. *Frontiers in Psychology*, *6*, 1–3.

Kahan, T. A., & Enns, J. T. (2010). Object trimming: When masking dots alter rather than replace target representations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 88-102.

Kahneman, D. (1967). An onset-onset law for one case of apparent motion and metacontrast. *Perception & Psychophysics*, *2*(12), 577–584.

Kahneman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, *70*(6), 404–25.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

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*(8), 2755–7.

Kinsbourne, M., & Warrington, E. (1962). The effect of an after-coming random pattern on the perception of brief visual stimuli. *Quarterly Journal of Experimental Psychology*, *14*(4), 223–234.

Kolers, P. A. (1962). Intensity and contour effects in visual masking. *Vision Research*, *2*(9–10), 277–294.

Kolers, P. A., & Rosner, B. S. (1960). On Visual Masking (Metacontrast): Dichoptic Observation. *The American Journal of Psychology*, *73*(1), 2.

Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *Journal of Neuroscience*, *27*(48), 13232–40.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494–501.

Lamme, V. A, Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, *9*(4), 406–13.

Lamme, V. A, & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–9.

Legge, G. E. (1978). Sustained and transient mechanisms in human vision: temporal and spatial properties. *Vision Research*, *18*(1), 69–81.

Lim, S. W. H., & Chua, F. K. (2008). Object substitution masking: When does mask preview work? *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1108-1115.

Liss, P. (1968). Does backward masking by visual noise stop stimulus processing? *Perception & Psychophysics*, *4*(6), 328–330.

Loffler, G., Gordon, G. E., Wilkinson, F., Goren, D., & Wilson, H. R. (2005). Configural masking of faces: evidence for high-level interactions in face perception. *Vision Research*, *45*(17), 2287–97.

Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: using apparent motion to isolate the object-level component of object substitution masking. *Journal of Experimental Psychology. Human Perception and Performance*, *29*(1), 106–20.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, *1*(2), 144–9.

Maffei, L., Cervetto, L., & Fiorentini, A. (1970). Transfer characteristics of excitation and inhibition in cat retinal ganglion cells. *Journal of Neurophysiology*, *33*(2), 276–84.

Matsumoto, N., Okada, M., Sugase-Miyamoto, Y., Yamane, S., & Kawano, K. (2004). Population dynamics of face-responsive neurons in the inferior temporal cortex. *Cerebral Cortex*, *15*(8), 1103-1112.

McDougall, W. (1904). The sensations excited by a single momentary stimulation of the eye. *British Journal of Psychology*, *1*(1), 78–113.

Mareschal, I., & Shapley, R. M. (2004). Effects of contrast and size on orientation discrimination. *Vision Research*, *44*(1), 57–67.

Merigan, W. H., Katz, L. M., & Maunsell, J. H. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *The Journal of Neuroscience*, *11*(4), 994–1001.

Merigan, W. H., & Maunsell, J. H. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, *16*, 369–402.

Müller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, *285*(5432), 1405-1408.

Monnier, M. (1952). Retinal, cortical and motor responses to photic stimulation in man. *Journal of Neurophysiology*, *15*, 469–486.

Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. Journal of Experimental Psychology. *Human Perception and Performance*, *31*(6), 1171–80.

Mumford, D. (1991). On the computational architecture of the neocortex. I. The role of the thalamo-cortical loop. *Biological Cybernetics*, *65*(2), 135–45.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics*, *66*(3), 241–51.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90–100.

Navon, D., & Purcell, D. (1981). Does integration produce masking or protect from it? *Perception*, *10*(1), 71–83.

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *Journal of Vision*, *15*(6), 4-4.

Öğmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Networks*, *6*(2), 245–273.

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

Phillips, W. A., & Singer, W. (1974). Function and interaction of on and off transients in vision I. Psychophysics. *Experimental Brain Research*, *19*(5), 493-506.

Piéron, H. (1925). I. Recherches expérimentales sur la marge de variation du temps de latence de la sensation lumineuse (par une méthode de masquage). *L'année psychologique*, *26*(1), 1-30.

Pilling, M., & Gellatly, A. (2010). Object substitution masking and the object updating hypothesis. *Psychonomic Bulletin & Review*, *17*(5), 737–42.

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*, 1560–1576.

Poggel, D. A., Treutwein, B., Calmanti, C., & Strasburger, H. (2006). Increasing the temporal g(r)ain: Double-pulse resolution is affected by the size of the attention focus. *Vision Research*, *46*, 2998–3008.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, *33*(7), 993-999.

Pollatsek, A., Rayner, K., & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 199–210.

Reicher, G. M. (1969). Perceptual recognition as a function of meaninfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275–280.

Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature*, *373*(6509), 66–68.

Reeves, A. (1982). Metacontrast U-shaped functions derive from two monotonic processes. *Perception*, *11*(4), 415–26.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature*, *387*(6630), 281–4.

Rogowitz, B. E. (1983). Spatial/temporal interactions: backward and forward metacontrast masking with sine-wave gratings. *Vision Research*, *23*(10), 1057–73.

Sackur, J. (2011). Dynamics of visual masking revealed by second-order metacontrast. *Journal of Vision*, *11*(4), 10.

Sally, S. L., & Gurnsey, R. (2004). Orientation discrimination across the visual field: matching perceived contrast near threshold. *Vision Research*, *44*(23), 2719–27.

Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. I. Review. *Psychologische Forschung*, *36*(1), 71–93.

Schultz, D. W., & Eriksen, C. W. (1977). Do noise masks terminate target processing? Memory and Cognition, 5(1), 90–96.Singer, W., & Creutzfeldt, O. D. (1970). Reciprocal lateral inhibition of on- and off-center neurones in the lateral geniculate body of the cat. *Experimental Brain Research*, *10*(3), 311–30.

Shooner, C., Tripathy, S. P., Bedell, H. E., & Öğmen, H. (2010). High-capacity, transient retention of direction-of-motion information for multiple moving objects. *Journal of Vision*, *10*(6):8, 1-20.

Singer, W., & Phillips, W. A. (1974). Function and interaction of on and off transients in vision II. Neurophysiology. *Experimental Brain Research*, *19*(5), 507-521.

Skottun, B. C. (2015). On the use of spatial frequency to isolate contributions from the magnocellular and parvocellular systems and the dorsal and ventral cortical streams. *Neuroscience and Biobehavioral Reviews*, *56*, 266–75.

Silverstein, D. N. (2015). A computational investigation of feedforward and feedback processing in metacontrast backward masking. *Frontiers in Psychology*, *6*(6), 1–14.

Spencer, T. J., & Shuntich, R. (1970). Evidence for an interruption theory of backward masking. *Journal of Experimental Psychology, 85*(2), 198–203.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors*, 5(1), 19–31.

Stewart, A. L., & Purcell, D. G. (1974). Visual backward masking by a flash of light: a study of u-shaped detection functions. *Journal of Experimental Psychology*, *103*(3), 553–66.

Stigler, R. (1910). Chronophotische Studien über den Umgebungskontrast. *Pflügers Archiv European Journal of Physiology*, *134*(6), 365-435.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*(6747), 869.

Tapia, E., & Breitmeyer, B. G. (2011). Visual consciousness revisited: magnocellular and parvocellular contributions to conscious and nonconscious vision. *Psychological Science*, *22*(7), 934–42.

Tapia, E., Breitmeyer, B. G., & Jacob, J. (2011). Metacontrast masking with texture-defined second-order stimuli. *Vision Research*, *51*(23-24), 2453–61.

Taylor, M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, *41*(4A), 782-787.

Tata, M. S., & Giaschi, D. E. (2004). Warning: Attending to a mask may be hazardous to your perception. *Psychonomic Bulletin & Review*, *11*(2), 262-268.

Turvey, M. T. (1973). On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, *80*(1), 1–52.

Uttal, W. R. (1970). On the physiological basis of masking with dotted visual noise. *Perception & Psychophysics*, *7*(6), 321–327.

Vidyasagar, T. R., Kulikowski, J. J., Lipnicki, D. M., & Dreher, B. (2002). Convergence of parvocellular and magnocellular information channels in the primary visual cortex of the macaque. *The European Journal of Neuroscience*, *16*(5), 945–56.

Weisstein, N. (1972). Metacontrast. *In Handbook of Sensory Physiology.* Vol. 7/4, *Visual Psychophysics* (ed D. Jameson and L.M. Hurvich). New York: Springer, pp. 233–272).

Weisstein, N., Ozog, G., & Szoc, R. (1975). A comparison and elaboration of two models of metacontrast. *Psychological Review*, *82*(5), 325–343.

Wiesstein, N. (1968). A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychological Review*, *75*(6), 494–521.

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, *23*(9), 873–82.

Wilson, H. R., Loffler, G., Wilkinson, F., & Thistlethwaite, W. A. (2001). An inverse oblique effect in human vision. *Vision Research*, *41*(14), 1749-1753.

Wilson, H. R., Loffler, G., & Wilkinson, F. (2002). Synthetic faces, face cubes, and the geometry of face space. *Vision Research*, *42*(27), 2909–23.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235.

Zhou, X., Chu, H., Li, X., & Zhan, Y. (2006). Center of mass attracts attention. *Neuroreport*, *17*(1), 85-88.