Encoding-decoding models of luminance contrast processing

by

Stuart Jackson

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

Center for Neural Science

New York University

January, 2016

Advisor: Wei Ji Ma

Acknowledgements

Thanks to all of my colleagues at the Center for Neural Science, particularly Wei Ji Ma and members of his lab, as well as colleagues from the labs of David Heeger and Marisa Carrasco. Also, special thanks to members of my defense committee for providing very useful feedback on an earlier version of this document.

A note about the thesis content: some of the work described here has been presented previously in abstract form, and may form the basis for future manuscripts. For information about any future related publications, or if you notice any errors in the current document, please feel free to get in touch (email: sj971@nyu.edu). Any future extensions, modifications or corrections to the work presented here will be documented separately. Thanks.

Abstract

There is a long tradition in sensory neuroscience of fitting precise neural computational models to experimental data. In a series of combined empirical and computational investigations, I illustrate important constraints on the encoding, retention, and read-out of information relating to luminance contrast in the visual world, a fundamental building block of vision. Using two-interval, forced-choice discrimination tasks, I first demonstrate that the efficiency of luminance contrast encoding-decoding is greatly impeded when high-contrast distractors appear in the opposite visual hemifield to a target stimulus; this behavior contrasts with relatively more efficient performance observed on an orthogonal task (orientation discrimination). I then explore a neural computational model of these results based on Fisher information, and find that, given a particular tuning parameterization, neither of two common models of sensory interaction satisfactorily explain both datasets simultaneously. In a later delayed-estimation experiment, I directly measure the precision with which single estimates of luminance contrast are encoded, maintained, and read-out from memory. The shape of observers' estimate distributions are adequately replicated by a probabilistic model of performance based on neurally-inspired components. In sum, the present thesis highlights key factors governing the precision of luminance contrast encoding and decoding, using complementary empirical and computational approaches. The thesis findings are also relevant to the broader literatures on attentional selection and the short-term retention of sensory information.

Contents

List of Figures

List of Tables

Perstando et praestando utilitati.

Original motto of New York University

1 Introduction

1.1 Overview

Progress in sensory neuroscience can greatly benefit from a tight interplay between experiment and model fitting. While some experimental work is exploratory or descriptive in nature, there is a long tradition of experiments explicitly designed in order to quantitatively distinguish theories. Successful quantitative model fitting allows for stronger conclusions about data than conventional model-free analyses, while deviations from model predictions or otherwise unexpected results provide strong clues as to how a theory must be modified.

The present thesis focuses on understanding the brain's ability to represent, maintain, and read-out information about the luminance contrast of stimuli in the world. As the fundamental building block for all of vision, information about light intensity passes through a variety of important neural processing stages: from retinal responses that signal individual spots of light, through sub-cortical and cortical processing stages, where a transformation towards representing mean luminance contrast in localized parts of the visual image is completed. Luminance contrast is arguably the most basic visual feature for pattern vision, and from it most other behaviorally-useful visual representations are derived (e.g., feature orientations, object boundaries, etc.). In this thesis, we present two sets of behavioral experiments with human observers focused on luminance contrast processing, as well as accompanying models based on contemporary theories of neural encoding and decoding.

Below, we first briefly review standard empirical approaches to understanding the nature of encoding and decoding for single visual stimuli, with a focus on luminance contrast processing. We then review key anatomical and functional characteristics of the visual system, focusing on how the visual system represents and transforms raw sensory input into the meaningful building blocks of vision, such as the amount of local luminance contrast or dominating orientation in small patches of the visual image. We then extend the discussion to more naturalistic conditions: emphasis is placed on the constraints governing behaviors that require encoding and retention of multiple stimuli over short intervals i.e., so called visual short-term memory (VSTM). The precision of encoding-decoding performance with topographically-structured neural representations (e.g., orientation-tuned neural responses) is contrasted with our current understanding of VSTM for luminance contrast. This debate, and the development of encoding-decoding models of luminance contrast processing,

forms the backbone of the thesis. On a more general note, we touch on the links between research on sensory encoding-decoding, VSTM, and the broader literature on attentional selection of sensory information. We conclude the introduction with a brief thesis synopsis.

1.2 Behavioral measures of stimulus encoding and decoding

Arguably the guiding pillars of modern computational approaches to understanding brain function are the inter-twined problems of neural encoding and decoding (Dayan & Abbott, 2001; Pouget et al., 2003). To maintain successful and productive behavioral repertoires, animals must adequately encode and make use of various sources of degraded information about the world. The development of realistic models of such behavior requires that researchers first characterize the everyday limits on encoding and decoding performance, using simplified experimental methods.

1.2.1 Definition of stimulus contrast

We begin by reviewing the standard behavioral paradigms used to study the processing of luminance contrast and other basic visual features. Stimuli in such tasks are commonly simplified spatial patterns such as sinusoidal gratings or circular discs, which allow for substantial experimental flexibility. Stimulus luminance contrast is typically given as a percentage of maximum contrast, and can be defined according to more than one convention, depending on the particular stimulus set-up and the experimenter's choice. For present purposes, when we refer to stimulus contrast, we define it according to Michelson contrast,

$$c_{\text{michelson}} = \frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}} \tag{1.1}$$

where L_{max} and L_{min} represent the maximum and minimum luminances of the stimulus respectively (e.g., the peak and trough of a sinusoidal grating). In Chapter 4, we make reference to an alternative definition (Weber contrast), which is described there.

1.2.2 Detection and discrimination of luminance contrast

Experimentalists have traditionally investigated the encoding-decoding of luminance contrast and other basic visual features using behavioral paradigms such as discrimination or detection (Nachmias & Sansbury, 1974; Legge & Foley, 1980; Skottun et al., 1987; Boynton et al., 1999). In a two-interval, forced-choice (2-IFC) discrimination task, for example, an observer is presented with a baseline or pedestal stimulus value across each of two intervals, and must correctly distinguish in which of the two intervals an additional increment (i.e., a contrast change) is added to the pedestal. Threshold performance is typically defined as the increment magnitude necessary to achieve some fixed performance criterion (e.g., 75% correct). A contrast-detection task is a limited form of discrimination, in which the baseline contrast is set to background luminance (i.e., 0% contrast), and absolute detection thresholds have commonly been collected alongside discrimination thresholds in the same experimental runs (Nachmias & Sansbury, 1974; Bradley & Ohzawa, 1986).

A classic finding in contrast-discrimination tasks is the improvement in discrimination performance for very low pedestal stimulus contrasts, relative to detection performance (Figure 1.1A) (Nachmias & Sansbury, 1974; Legge & Foley, 1980; Bradley & Ohzawa, 1986). This effect is thought to arise from an early accelerating non-linearity in contrast encoding, under the assumption that regardless of stimulus contrast level, a fixed change in some internal response (i.e., neural firing) is required for discrimination of changes to the stimu-

lus (Nachmias & Sansbury, 1974; Legge & Foley, 1980; Bradley & Ohzawa, 1986). Following this early facilitation effect, a gradual increase in thresholds for much of the stimulus contrast axis is typically observed, with the slope of this increase often found to be around 0.5-0.7 on a log-log axis (Nachmias & Sansbury, 1974; Legge & Foley, 1980). Presumably, the gradual increase in threshold with increasing pedestal stimulus contrast reflects some trade-off between the shape of the internal response to stimuli and internal noise levels, a debate which has continued for some time (Gorea & Sagi, 2001). In passing, however, we note that a number of studies which have measured thresholds for very high pedestal contrasts (e.g., above 50% contrast) have found some late flattening or decrease in threshold, thereby suggesting that the later part of the contrast-discrimination function is not necessarily monotonic throughout (Kingdom & Whittle, 1996; Zenger-Landolt & Heeger, 2003; Chirimuuta & Tolhurst, 2005; Pestilli et al., 2011).

1.2.3 Alternate measures of luminance contrast processing

A number of studies have also utilized a matching or adjustment protocol to study luminance contrast processing (Georgeson & Sullivan, 1975; Prinzmetal et al., 1997). For example, Georgeson & Sullivan (1975) had observers adjust the contrast of one of two sinusoidal gratings presented side-by-side, so as to match the other grating in contrast: for relatively broad differences in stimulus spatial frequency, observers could accurately match the contrasts of the variable and standard stimuli, suggesting a substantial degree of adaptability on the part of the local neural mechanisms feeding into stimulus contrast coding. In addition, the effect of luminance contrast on the encoding of another important visual feature, stimulus orientation, has also been studied using discrimination protocols (Blake & Holopi-

gian, 1985; Skottun et al., 1987; Mareschal & Shapley, 2004). As stimulus contrast increases, a characteristic decrease in thresholds for orientation discrimination is found, with typically no substantial improvement in performance beyond stimulus contrasts of about 10-20% (Figure 1.1B). In other words, observers appear to reach ceiling performance levels for orientation discrimination at relatively low-to-moderate contrasts, with the bound on performance likely set by fixed levels of internal noise (Mareschal & Shapley, 2004).

1.3 Neural basis of stimulus encoding and decoding

To better understand the foundation on which behaviors such as contrast discrimination are based, we now briefly describe the architecture and function of the early visual system, and the neural representations that act as the building blocks of cortical vision. We also briefly discuss the known characteristics and effects of neural noise on sensory processing.

1.3.1 Architecture and function of the early visual system

Anatomical and physiological studies in numerous species illustrate common neural processing architectures for visual information (McIlwain, 1996). The primate visual system, for example, is known to progress along two main processing streams which differ in their anatomical and physiological characteristics, the magnocellular and parvocellular pathways (Shapley, 1990; McIlwain, 1996). This pathway segregation begins in the retina, and becomes highly evident in the layering of the major sub-cortical visual relay, the lateral geniculate nucleus (LGN) of the thalamus. The two pathways then converge onto different sub-layers of the primary visual cortex or V1, with magnocellular cells terminating in layer 4Cα and with a large portion of the parvocellular pathway cells terminating in layer

4Cβ (McIlwain, 1996; Sincich & Horton, 2005). From here, the organization of the separate processing streams becomes more nuanced in layout, with the parallel processing that predominated at earlier synapses (e.g., retina to thalamus), giving way to cross-talk in cortex. For example, V1 neurons receiving thalamic inputs may have subsequent synapses onto other layers of V1, and individual cortical layers can be reciprocally connected to one another (Callaway, 2003; Sincich & Horton, 2005).

Broad functional differences are apparent across these and other parallel processing channels in the early visual system. For example, the transmission of chromatic (i.e., color-related) information from retina to cortex is subserved primarily by parvocellular pathway processing, while the magnocellular pathway is thought to play a more dominant role in processing achromatic, luminance-defined signals and motion (Lennie et al., 1990; Shapley, 1990; Johnson et al., 2001). In fact, cells in the respective pathways illustrate broadly divergent sensitivity to luminance intensity and contrast, presumably playing qualitatively distinct roles in luminance contrast processing (Kaplan & Shapley, 1986; Shapley, 1990). In parallel, from the retinal bipolar layer onward, sensory responses are carried by cells that become either more or less responsive when stimulated within the center of their receptive field by light increments, the so-called ON/OFF channels in vision (Wiesel & Hubel, 1966; Schiller et al., 1986; McIlwain, 1996). Presumably, this early separation of responses to increments and decrements in light evolved to allow for maximum system contrast sensitivity at minimal biophysical cost (Schiller et al., 1986). Of note, numerous findings in recent years have now definitively illustrated asymmetries in the early representation of positive and negative luminance signals (Chubb et al., 2004; Yeh et al., 2009; Ratliff et al., 2010; Kremkow et al., 2014), a topic we will briefly touch on in later discussion.

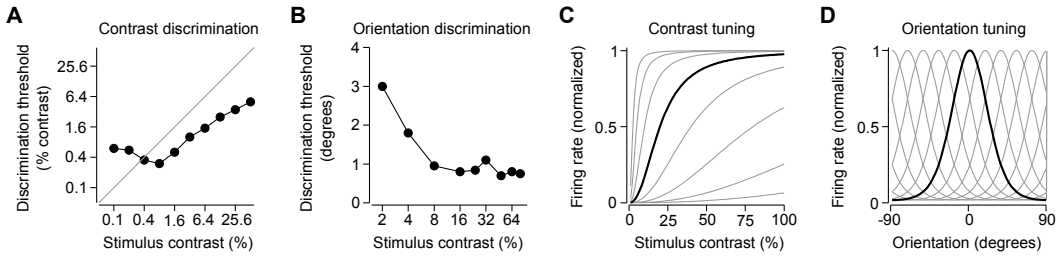

Figure 1.1: Discrimination behavior and tuning functions for contrast and orientation. A) Contrast-discrimination performance for sinusoidal gratings of increasing contrast, adapted by eye from Legge & Foley (1980). B) Typical orientation-discrimination performance for grating stimuli of increasing contrast, adapted by eye from Skottun et al. (1987). C) Idealized (Naka-Rushton) single-neuron contrast response functions with logarithmically-spaced semi-saturation constant. D) Idealized (Von Mises) orientation tuning curves tiling the orientation axis.

1.3.2 The building blocks of cortical vision

A series of transformations occur in early vision from representing local, absolute light intensity values in neural responses, to representing luminance contrast in the scene being viewed. The visual system achieves this through a cascade of gain control processes, where the input drive at a given stage of processing (e.g., V1) is scaled by a broader, suppressive (i.e., inhibitory) signal in the surrounding neural circuit (Shapley & Victor, 1979; Ohzawa et al., 1985; Wilson, 1999; Carandini & Heeger, 2012). Gain control processes serve to optimize the visual system's response under varying input conditions (e.g., for a given time of day or mean luminance level), by ensuring that the system is most sensitive to fluctuations around the mean input level. In general, increases in stimulating contrast typically produce a monotonically increasing contrast response profile, with lower contrasts producing relatively fewer mean spikes, and higher contrasts eventually saturating the response of the neuron (Figure 1.1C). This firing rate behavior can be captured by relatively simple mathematical expressions, such as the commonly used Naka-Rushton equation (Naka & Rushton, 1966; Albrecht & Hamilton, 1982). By describing the responses of individual neurons using

simplified equations with minimal parameters, properties of larger groups of neurons (i.e., local populations) can be summarized in terms of distributions of parameter values (Albrecht & Hamilton, 1982). As stimulus contrast is encoded in a neuron by the strength of its firing, we refer to luminance contrast throughout this thesis as being an intensity-coded feature.

Unlike the intensity-coding used for luminance contrast processing, the detailed architecture of the early visual system provides the ideal substrate for the structured or topographic representation of other basic visual features. In now classic experimental work, Hubel & Wiesel (1962) demonstrated topographic representation along several visual stimulus feature dimensions: e.g., for location (retinotopy), orientation (orientation pinwheels) and disparity processing (ocular dominance columns). In the intervening decades, neuroscientists have attempted to further refine our understanding of these topographic representational maps, perhaps most successfully for orientation-tuned neural responses (Ferster, 2003). In contrast to the monotonic tuning functions found for contrast, local orientations in an image are mapped onto approximately symmetric neural tuning functions for orientation, with individual neurons typically responding to a limited range of preferred stimulus orientations. An idealized example of symmetric, orientation tuning curves is depicted in Figure 1.1D.

1.3.3 Neural noise and the encoding-decoding process

Even with a detailed supporting neural architecture, performance in behavioral tasks is never perfect: noise accrues in an observer's representation of the visual world, from the initial photon transduction process in retina, through to noisy neural spiking at sub-cortical

and cortical layers of the visual system. Numerous investigators have studied characteristics of neural noise in visual cortex (Tolhurst et al., 1981; Shadlen & Newsome, 1998; Goris et al., 2014). A key finding relates to the link between the mean and variance of a neuron's firing rate: these have often been found to scale together with a ratio of approximately 1, meaning that a neuron's spiking behavior can be reasonably well-approximated as a Poisson process (Tolhurst et al., 1981; Shadlen & Newsome, 1998). However, this relationship is certainly not exact for real neurons, and the ratio of variance-to-mean (i.e., the Fano factor) has been found to range above and below 1 (Shadlen & Newsome, 1998; Goris et al., 2014). In fact, an appreciation has grown recently for the importance of accounting for trial-to-trial gain fluctuations in models of firing rate statistics (Goris et al., 2014). Fluctuations in the strength of neural response to a given stimulus might arise from a variety of non-sensory sources e.g., through modulations of attention or arousal. Models of encoding-decoding performance have begun to incorporate more realistic gain fluctuations into their computational architecture (May & Solomon, 2015), a topic we will touch on briefly in later discussion.

1.4 Encoding-decoding performance under more natural conditions

The distinctions between research on basic stimulus encoding-decoding, VSTM, and attentional selection are sometimes subtle. While VSTM paradigms inevitably incorporate delay between the time of encoding and read-out, this distinction is not typically considered in standard 2-IFC discrimination tasks, where stimuli must still be held in memory from one interval to the next. To better link these topics, we first describe how temporal delay affects the classic discrimination behaviors described earlier. We then flesh out our current under-

standing of how encoding-decoding performance changes under more natural conditions, such as in the presence of target location uncertainty (i.e., for multiple stimulus displays). This discussion is divided into two parts: research that has investigated the effects of varying set-size on performance, and research that has investigated the effects of stimulus interactions on performance. In doing so, we introduce a variety of research on VSTM, with particular emphasis on the delayed-estimation paradigm. We also touch on the broader topic of attentional selection of sensory information.

1.4.1 Role of temporal delay on encoding-decoding performance

The similarities and differences between perception and VSTM are immediately apparent when one compares VSTM performance for luminance contrast to performance for features such as stimulus orientation. Contrast-discrimination thresholds increase substantially with response delay periods longer than a few seconds, unlike performance profiles for stimulus properties such as orientation, motion, or spatial frequency, which are much less affected across prolonged delays e.g., 10 s or more (Magnussen & Greenlee, 1999). These findings likely reflect differences in how the relevant stimulus information is neurally represented (i.e., intensity vs. topographic encoding), and how these representation decay over time. In the case of orientation, the topographically-arranged orientation maps provide a natural substrate for precise encoding, while no analogous representation facilitates the precise encoding of contrast.

1.4.2 Role of set-size on encoding-decoding performance

The effect of stimulus set-size on encoding-decoding precision has been extensively studied, both for attention-related tasks such as visual search (Palmer et al., 2000; Mazyar et al., 2012) and for studies of VSTM using paradigms such as change detection and change localization (Pashler, 1988; Luck & Vogel, 1997; Bays & Husain, 2008). In earlier studies, data were often interpreted as reflecting a bound on performance as a function of set-size, with set-sizes greater than four thought to require more than the available number of encoding slots in memory (Luck & Vogel, 1997; Cowan, 2001). This view has changed in recent years. For example, visual search performance has now definitively been shown to systematically decrease as the number of stimuli that make up the search array increases (Mazyar et al., 2012), and similar effects hold for VSTM paradigms (Bays & Husain, 2008; Ma et al., 2014).

Researchers have recently taken a more direct approach to measuring the precision of stimulus encoding and decoding, by utilizing the delayed estimation method (Wilken & Ma, 2004; Fougnie et al., 2012; van den Berg et al., 2012), a paradigm inspired by earlier matching experiments (Prinzmetal et al., 1997, 1998). In a delayed-estimation task, an observer is first presented with a brief stimulus display, and after a short delay, must attempt to match the input stimulus precisely. By systematically controlling delay time (typically on the order of a couple of seconds), such tasks serve to focus on the processes of encoding and maintenance of information in VSTM. For example, in a subset of their experiments, Wilken & Ma (2004) had observers estimate the orientation, color, or spatial frequency of a stimulus presented 1500 ms earlier. As set-size increased, estimate precision decreased in monotonic fashion, counter to the traditional view of VSTM as being made up of slots (Luck & Vogel, 1997; Cowan, 2001). By providing a more continuous measurement of en-

coding quality, such tasks provide richer datasets that lend themselves more naturally to detailed computational model-fitting.

1.4.3 Role of stimulus interactions on encoding-decoding performance

The effects of distractor stimuli on encoding-decoding precision goes beyond the basic set-size effect. Even when an observer knows in advance that only one stimulus is relevant for a task, there can still be an effect of irrelevant distractors on performance. Systematic investigation of such effects has so far been limited; yet, a number of recent investigations on the topics of attentional selection and VSTM have utilized paradigms where the effects of direct stimulus interactions on behavioral performance can be studied. Of these, a number of studies relevant to the topic of luminance contrast encoding-decoding have emerged (Pestilli et al., 2011; Chen & Seidemann, 2012; Hara & Gardner, 2014; Itthipuripat et al., 2014), as well as empirical and theoretical investigations studying the general effects of feature and stimulus interactions on decoding (Bays et al., 2009; Matthey et al., 2015; Orhan & Ma, 2015).

Results of some of these studies support the view that VSTM for luminance contrast is severely hampered by the presence of irrelevant distractors. This behavior was nicely demonstrated in a series of contrast-discrimination experiments in which observers had to either focus or distribute their attention across multi-item displays (Pestilli et al., 2011; Hara & Gardner, 2014). The presence of a single high-contrast distractor was sufficient to severely disrupt threshold performance at the target location (Pestilli et al., 2011). The large responses evoked by high-contrast distractors must have dominated in the selection of sensory signals for decision, supporting a model of encoding-decoding using something akin

to a max-pooling operation as a decision rule (Pelli, 1985; Palmer et al., 2000; Pestilli et al., 2011).

Other recent findings provide a window into the question of stimulus interactions and encoding-decoding precision. For example, using a delayed-estimation paradigm, Bays et al. (2009) found that estimates of the color of a recently presented target item appeared to be accompanied by a sizeable proportion of non-target color reports. The authors described a model of this process in which memory for item color and location could interact, thereby systematically affecting the shape of estimate response distributions for target color. This type of effect highlights the difficulty faced by the brain in disentangling the actual sources of sensory responses in the context of multiple stimulus displays. Presumably, as has been found at numerous levels of visual processing, individual neurons might compute weighted sums of constituent inputs, effectively mixing neural responses originating from separate sources (Recanzone et al., 1997; Zoccolan et al., 2007; Busse et al., 2009). This line of thinking has recently been formalized in an encoding-decoding architecture based on the linear-mixing of separate sensory neural responses (Orhan & Ma, 2015). In Chapter 3, we explore a model of this sort.

1.5 Thesis synopsis

The present thesis consists of a complementary pair of investigations that further our understanding of the neural representations and computational rules governing the encoding and decoding of luminance contrast. Each investigation was comprised of multiple psychophysical experiments with human observers, and involved subsequent fitting of neural computational models to each dataset. The thesis builds on a wide variety of studies that

relate to the topics of sensory encoding and decoding.

In Chapter 2, we present results from a series of 2-IFC discrimination experiments that directly question the principled nature of VSTM for luminance contrast: in the presence of irrelevant distractor stimuli, contrast-discrimination performance is found to deteriorate more substantially relative to decoding performance for another low-level, visual feature (orientation). We investigate these differences in Chapter 3, by simultaneously fitting particular neural response models to each of the datasets described in Chapter 2. Neither of two common models of sensory interaction (i.e., divisive normalization, linear mixing of neural responses) provides a satisfactory explanation of both datasets simultaneously; we discuss possible reasons for the failure of these models to account for performance. We then dig deeper in Chapter 4, studying the trial-by-trial precision of VSTM for luminance contrast. Results are presented from a series of delayed-estimation experiments, allowing us to more clearly define the likely neural constraints governing the encoding and short-term retention of luminance contrast information. Behavior of human observers on these tasks is shown to be highly principled: performance is captured satisfactorily by a probabilistic model of neural responses incorporating biologically-plausible model components.

2

Efficient and inefficient selection from the same sensory neural response: psychophysics

2.1 Introduction

To successfully perform even basic visual tasks, humans must efficiently select and manipulate relevant environmental information. The brain processes this sensory information in various stages, from initial encoding, to the maintenance of signals in visual short-term memory, to the application of appropriate sensory read-out rules. A long tra-

dition of research in sensory neuroscience has studied such processing using very simple stimuli and tasks, allowing for maximal experimental control and model tractability. For example, human visual performance has been commonly characterized by studying how observers discriminate the contrasts (Nachmias & Sansbury, 1974; Legge & Foley, 1980; Boynton et al., 1999) or orientations (Westheimer et al., 1976; Blake & Holopigian, 1985; Skottun et al., 1987) of two successively presented sinusoidal gratings, so-called two-interval, forced-choice (2-IFC) discrimination tasks. Performance or threshold on such tasks is typically defined as the amount of stimulus change needed to achieve a criterion level of behavioral performance (e.g., 75% correct). When plotted as a function of the pedestal stimulus contrast, contrast-discrimination thresholds typically follow a 'dipper' shape, first decreasing with small increases in pedestal contrast above background luminance, and then increasing at higher values of pedestal contrast (Nachmias & Sansbury, 1974; Legge & Foley, 1980). On the other hand, as pedestal contrast increases, the deviations in pattern orientation necessary to reach a threshold level of orientation-discrimination performance typically decrease in a systematic and monotonic fashion (Skottun et al., 1987).

Single-stimulus tasks afford the researcher precise experimental control and simplicity, yet such paradigms are entirely unlike natural vision, in which multiple objects are present at once. In fact, despite substantial progress in understanding how discrimination behavior is linked to local neural computations (Paradiso, 1988; Boynton et al., 1999; Sanborn & Dayan, 2011; Berens et al., 2012), relatively little is known about the selection or decoding strategies implemented by observers when faced with more complex sensory input (i.e., multiple possible target stimuli). Thus, there is a need to develop experimental paradigms where the joint encoding of multiple stimuli is required for successful performance. We

focus here on the role of irrelevant distractors in multiple stimulus encoding; specifically, on tasks in which the observer attempts to select from memory information from a single, post-cued stimulus location (Pestilli et al., 2011; Sergent et al., 2011; Hara & Gardner, 2014; Itthipuripat et al., 2014). We do not consider tasks in which all items are actually relevant for the task, such as global target detection (Palmer et al., 2000; Ma et al., 2011).

How do irrelevant distractors influence behavioral performance in traditional discrimination tasks? While direct evidence is so far weak, there are hints that discrimination performance with more complex sensory input varies as a function of the particular task and encoding constraints under investigation. For example, unlike the known topographic neural representation for stimulus features such as orientation, the neural representation for contrast is based fundamentally around response intensity (Albrecht & Hamilton, 1982), a characteristic that likely hinders the formation of abstract memory representations for contrast (Xing et al., 2014). Thus, sensory evidence for a particular contrast may be available only briefly after stimulus disappearance, and in imprecise form, encouraging non-selective pooling across stimuli when multiple estimates are made simultaneously (Pestilli et al., 2011; Hara & Gardner, 2014). For example, Pestilli et al. (2011) found that when attention was distributed across multiple stimuli that varied in contrast, the presence of a single high-contrast distractor was sufficient to severely disrupt contrast-discrimination performance at a target location. However, sensory evidence for other stimulus properties (e.g., orientation) may be maintained with greater precision and for longer post-stimulus delays (Magnussen & Greenlee, 1999). Thus, when selecting from sensory signals at different spatial locations, performance on tasks such as orientation discrimination may be much less influenced by distractors or prone to sub-optimal decision rules, a conjecture with some

experimental support (Sergent et al., 2011).

To test this possibility, we ran separate experiments in which observers discriminated changes to either the contrast or orientation of a target stimulus presented in the hemifield opposite to a distractor. Identical stimulus protocols were used across experiments, the only difference being that small contrast increments were added to the target location in the contrast experiment, while small orientation deviations were added to the target in the orientation experiment (Figure 2.1). To measure the effect of distractor strength on performance, we systematically varied the pedestal contrasts assigned to targets and distractors. Results reinforced the view that for the two forms of discrimination tested, observers appear to select from identical sensory neural responses in incommensurate ways. In a subsequent chapter, we will attempt to fit these data using a precise neural computational model.

2.2 Materials and Methods

2.2.1 Participants

Data from the same eight observers (two authors) were collected in both of the main experiments. Five observers completed the contrast-discrimination experiment prior to the orientation-discrimination experiment, three observers ran in reverse order. Experimental sessions were typically performed over a 1-3 week period, with the different experiments separated by up to several months. Observers were recruited from the general student/staff body at New York University (paid $10/hr) and amongst lab colleagues, and had varying degrees of experience in psychophysical testing. One additional recruit was not tested beyond the practice session, during which this individual confirmed being diagnosed with an attention-related disorder. Aside from the authors, observers had no knowledge of the

specific experimental hypotheses. All observers gave written informed consent, and experiments were carried out with approval of the NYU University Committee on Activities Involving Human Subjects.

2.2.2 Contrast discrimination

Task

We tested the effects of distractors of different contrast on contrast-discrimination performance at a post-cued target location (Figure 2.1). Each trial began with the presentation of pre-stimulus arrows pointing left and right of fixation (1 s), cueing the observer to distribute attention equally to two peripheral locations (6° eccentricity), while remaining fixated on a central fixation cross (1° width). After a short delay (100 ms) and auditory tone indicating stimulus onset, a pair of gratings were briefly presented in a first stimulus interval (600 ms), one positioned left and the other right of fixation along the horizontal meridian. The gratings then disappeared, and after a short ISI (200 ms), the gratings reappeared for a second stimulus interval (600 ms). A positive contrast increment was added to one of the gratings (the 'target') in one of the two intervals. After the stimuli had left the screen, there was a second short delay (400 ms), followed by presentation of a green arrow indicating the target location. Observers responded during this interval by pressing one of two keys on the keyboard ('1' or '2'), judging which of the two stimulus intervals contained a higher contrast at the target location. This response interval was of fixed duration (1200 ms), and trials with no response were not replaced. Observers received feedback on each trial (color change of the fixation cross and auditory tone), and were instructed to perform as accurately as possible throughout sessions. Trials were separated by an ITI of pseudo-random duration

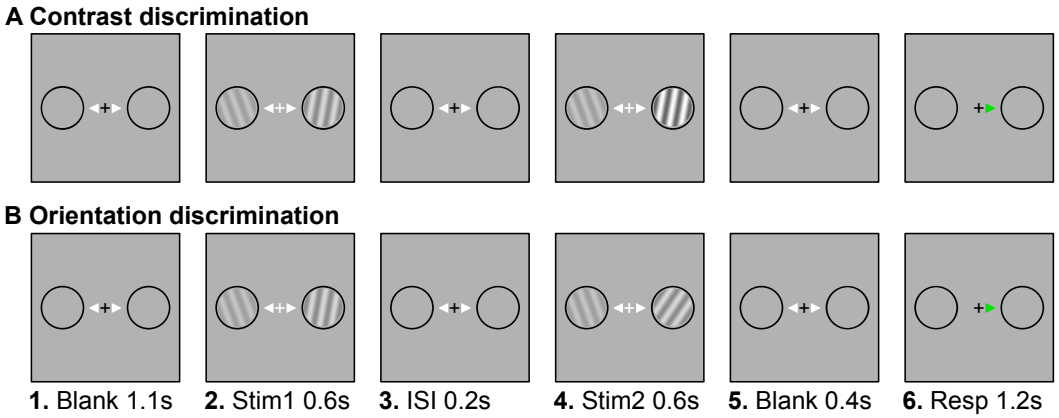

Figure 2.1: Sensory discrimination in the presence of distractors: experiments. Contrast- and orientation-discrimination experiments were carried out in separate sessions (i.e., non-interleaved), using a largely identical set-up. Trials began with the presentation of pre-cues pointing left and right of fixation, cueing the observer to distribute attention equally to two peripheral locations (6° eccentricity). A pair of gratings (5° diameter) were then briefly presented in the first stimulus interval, left and right of fixation. The gratings then disappeared, and after a short ISI, reappeared for a second interval. In the contrast-discrimination experiment, a positive contrast increment was added to a target grating in one of the two intervals; in the orientation-discrimination experiment, a clockwise or counter-clockwise orientation increment was added to the target grating. After the stimuli had left the screen, a post-cue (green arrow) indicated the target location. Observers judged in which of the two intervals the target grating had either higher contrast or was rotated more clockwise.

(800-1200 ms, 100 ms steps). Observers completed five sessions (480 trials per session, approximately 1 hr duration). We regarded the first session as a practice session and analyzed only the final four sessions. Observers received a mandatory rest period after every block of 120 trials, and could also pause presentation at any time by pressing the space bar.

Stimulus design and experimental conditions

Stimuli were presented in a darkened room on a gamma-corrected CRT (75 Hz, 1152 x 870 resolution), and were generated using MATLAB (The Mathworks) and MGL (see www.justingardner.net/mgl). For the first phase of experiments, we used a gamma correction table calculated some time prior to testing (approximately 24 months earlier); a later correction carried out prior to the control experiment suggested that drift had occurred

from the original table values. However, mean and maximum luminances always lay in typical ranges for these types of experiments (approximately 35 and 70 cd/m^2 respectively, plus or minus a small amount). We are not concerned that this drift had any meaningful consequences for our stimulus comparisons, given that we used a relatively coarse sampling of contrasts from the entire luminance contrast range. Gratings were counter-phase flickering sinusoids (5 Hz, 2 cycles/°) measuring 5° in diameter. Gratings were presented inside black circular frames, such that a small gap lay between the frame and the grating edge (raised-cosine, edge width 0.5°). Target and distractor gratings could appear with one of four pedestal contrast values (10, 20, 40 or 80% Michelson contrast). An experimental condition is defined as a combination of target and distractor contrast. All combinations of target and distractor contrast were presented, excluding those conditions where target and distractor would appear with identical pedestal contrast. The reason for this exclusion was that in such conditions, an ideal observer could use information from a single interval to perform the task, as positive contrast increments were always added to the target pedestal. Thus, we measured contrast-discrimination thresholds for twelve target-distractor contrast pairs in total.

Forty trials per condition were presented in each session, and conditions were randomly interleaved. Of the forty trials per condition, increments on thirty-two trials were controlled by an adaptive, 1-up-2-down staircase (i.e., the increment was increased after an incorrect response, and decreased after two consecutive correct responses). Increments on the remaining trials were hand-picked on a session-by-session basis. For the hand-picked increment trials, which were randomly interleaved with staircase trials, increments were typically set to low and/or high values so as to improve the quality of fit to baseline/asymptotic

performance across session (occasionally some intermediate fixed increments were also presented). In all conditions except for 40% and 80% pedestal contrasts, we used contrast increments of 0.5, 1, 2, 4, 8, 12, 16, 24, 32, 48 and 64%. When pedestal contrast was 40% or 80%, we modified the increment array such that contrast would not exceed 100% (e.g., for an 80% contrast target, the largest possible increment was 20%). In a couple of early sessions, we included increments of 6, 10, 20 and 40% in the general array; we retained any trials using these increments for analysis, except for a small handful (eight trials for one observer) in which target-plus-increment inadvertently equaled distractor contrast. Staircase endpoints from the practice session were used as staircase starting points in the first test session, and similarly each subsequent test session began with staircase endpoints from the session prior. In anticipation of the orientation-discrimination experiment, target and distractor gratings took one of ten pseudo-random orientations (9, 27, 45, 63, 81, 99, 117, 135, 153 or 171°). For each of the twelve conditions present within a 120-trial block, each of the ten possible orientations was used as target orientation exactly once. The frequency of distractor orientations was not controlled in a similar fashion, but pseudo-randomized such that distractor orientation always differed from target orientation on any given trial. On any trial, orientations were held constant across intervals.

2.2.3 Orientation discrimination

Task, stimulus design and experimental conditions

Aside from a separate training phase (see below) and the increment type (i.e., orientation), all aspects of stimulus presentation and protocol were largely identical to the contrast-discrimination experiment. On each trial, a clockwise or counter-clockwise orientation in-

crement was added to the target grating, and observers judged which interval contained the more clockwise-oriented stimulus at the target location, which was post-cued as in the contrast task. In addition, all sixteen combinations of target-distractor contrast were presented (all combinations of 10, 20, 40 and 80% contrast). Given the larger number of conditions, observers completed six 480-trial sessions, with thresholds estimated from the final five sessions. Each session was divided into three 160-trial blocks, and given the slightly longer blocks, observers were encouraged to pause presentation once or twice per block as needed. In each session, thirty trials were presented for each of the sixteen target-distractor contrast conditions. The staircase approach was similar to that used for the contrast-discrimination experiment, with a fixed array of orientation increments provided for each staircase (increments of 1, 2, 4, 6, 8, 12, 16 and 24°), and with a small proportion of trials set to increments which were hand-picked from the fixed array on a session-by-session basis.

Training protocol

Our training protocol was informed by an initial version of the orientation-discrimination experiment which included 32° orientation increments, and in which two observers completed several sessions each (a practice session plus two/three test sessions, respectively). For both observers, we found that performance in a large number of conditions was near chance even for the 32° increments. This may have resulted from several factors, including the large range of pedestal orientations used (spanning 180°), the time-limited response interval, and the direction ambiguity inherent in circular orientation space (which might be particularly problematic for very large orientation increments).

We excluded the data from these early sessions from further analysis and in response to

the noted problem, made several modifications to the experimental design. First, we limited the orientation increments in the main experiment to a maximum of 24°. Second, we introduced a training protocol with fixed, large orientation increments (24°) only. During training, observers were verbally encouraged to focus on the global, rotational nature of the discrimination (as opposed to focusing on the left/right tilt of one end of the grating), and were instructed that the large orientation increments presented in training were the largest possible orientation changes they would experience throughout testing; feedback from several observers confirmed that the large increments were typically well above detection threshold. In an effort to step up the difficulty level of training blocks gradually, we used three different block types that differed in complexity: the first involved orientation discrimination for a single grating presented over two intervals; the second used two-grating displays, and involved orientation discrimination for a pre-cued target grating; the third used two-grating displays but without target pre-cueing, as in the main experiment. Blocks consisted of forty trials each. For the single-grating blocks, stimuli were presented at fixation and each of the four test contrasts were presented an equal number of times, in randomly interleaved fashion. For the two-grating blocks, we used pseudo-randomly selected subsets of the contrast conditions from the main experiment. Specifically, we ensured in each block that each of the four contrasts appeared once as target and once as distractor, and that conditions were matched in opposite pairs so that the target could not be distinguished based on contrast. Observers were informed of their percentage correct after each block, and moved from one block type to the next when they scored consistently in the range of 80-90% and above (this was judged online by the experimenter on an observer-by-observer basis, as observers differed noticeably in the consistency with which they could

score above 90% for these relatively simplified blocks). In total, the different observers completed from one to three training sessions of approximately 1 hr each. In a final modification of the original protocol, observers also completed one or more short warm-up blocks at the beginning of each test session. Again, the exact number of blocks was determined online by the experimenter, based on a rough appraisal of the consistency of the observer's performance relative to the pre-test training sessions.

2.2.4 Threshold estimation and statistical tests

Main analyses

We estimated thresholds separately for each observer in each experiment. For a given experiment, we combined data from the final four (contrast) or five (orientation) sessions, so that we had a maximum of 160 (contrast) or 150 (orientation) trials per condition from which to estimate threshold. Trials on which no response was made were excluded from analysis (less than 1% of trials), meaning that the actual numbers of trials per condition were on average slightly less than the numbers above. We fit a Weibull function to the data from each condition,

$$\text{Proportion correct}(\Delta s) = 0.5 + \left(0.5 - p_{\text{lapse}}\right)\left(1 - \exp\left(-(\frac{\Delta s}{a})^{b}\right)\right), \qquad (2.1)$$

where Δs is the orientation or contrast increment, $1 - p_{\text{lapse}}$ is the asymptotic proportion correct at very large increments, and the parameters a and b control the midpoint (i.e., bias) and steepness (i.e., slope) of the psychometric function, respectively. To impose the constraint of a fixed lapse rate across conditions, we adopted the following fitting procedure.

We varied p_{lapse} in steps of 0.001 between 0.001 to 0.06. At each step, we used maximum-likelihood estimation, implemented using the MATLAB function *fminsearch.m*, to find the best-fitting a and b independently for each condition. The log likelihoods of the resulting fits to the individual conditions were then summed, to give a combined log likelihood value for each lapse rate. The p_{lapse} with the highest log likelihood, and the associated collection of bias and slope value pairs, were taken as the best-fitting parameter values for an individual observer. We then derived thresholds from the fitted psychometric functions by finding the increment necessary to achieve 75% correct performance (Figure 2.2A). A small number of occurrences of threshold going outside increment boundaries were left as is (e.g., contrast-discrimination threshold exceeding 20% for 80% contrast targets); they played no meaningful role in comparisons of most interest.

To measure the probability of observing mean threshold differences by chance for different distractor contrast levels, the following statistical test was run. Using a randomization analysis, we calculated p-values for each possible distractor pair comparison, separately at each target contrast value. For each comparison, this involved pooling the pair of distractor response distributions for an observer, and re-sampling trials for the two conditions from the randomly shuffled pooled distribution. Psychometric functions were fit to the re-sampled pair, and threshold difference computed (with p_{lapse} held fixed at the value calculated in the main analysis). This sampling procedure was repeated 10,000 times to generate a distribution of threshold differences for each pair. The same process was applied to each observer's data separately, such that for each comparison we had eight observers x 10,000 threshold differences. From these, we calculated a mean distribution for each comparison (averaging across the observers' unsorted distributions at each sample). The resulting mean

distribution was sorted from smallest to largest, and the probability of observing by chance the measured mean threshold difference was read off from this sorted distribution (by finding the index with minimum absolute difference from the measured mean threshold difference, ignoring the sign of that minimum difference). Exact p-values are reported in text and in Table 2.1 (to four decimal places, and with a minimum possible value of $p = 0.0001$).

Control analyses

As a precautionary measure, we repeated the main analyses with two modifications. First, we re-ran the randomization analysis after excluding trials containing saccades during the stimulus intervals (see Eye-movement recording). Second, to account for the possibility that some fits might have been over or under-dispersed due to, for example, across-session learning or experimenter bias in handpicking increment values, we re-calculated the p-values after performing a deviance analysis on the psychometric function fits (Wichmann & Hill, 2001). Specifically, the deviance statistic was computed for each psychometric function fit, and then compared to a distribution of simulated deviance statistics (10,000 samples). This distribution was computed using the original fit parameters as generating model, and calculating deviance at each sample between the simulated data and the best-fit psychometric function to the simulated data. Individual fits whose deviance statistic lay outside a relatively narrow confidence interval (84%) were excluded from p-value calculation.

We also carried out two additional control analyses on sub-portions of the data. First, we determined whether the increased exposure observers had from pre-training on the orientation task dampened distractor effects in subsequent test sessions, relative to the size of ef-

fects observed in the contrast-discrimination task at least. To test this possibility, we pooled individual observer data from the first two test sessions of the orientation-discrimination task (excluding the practice run), to create a super-observer dataset from which we estimated thresholds. We created a similar pooled dataset from the contrast-discrimination experiment, this time combining data from the final two test sessions, so that our comparison was between sessions in which observers had approximately as much or more prior exposure in the contrast-discrimination task. A randomization analysis was performed on psychometric function fits to the pooled datasets.

In a final control analysis, we determined whether the difference in orientation between target and distractor stimuli played any role in performance in either experiment (e.g., through grouping of similar orientations, or other spatially broad interactions). To test this possibility, we divided the datasets into two parts - in one part, we placed trials in which the distractor was oriented either 18° or 36° clockwise or counter-clockwise of the target (i.e., closer to parallel); in the other, we placed trials in which the distractor was oriented 54°, 72° or 90° away from the target (i.e., closer to orthogonal). It was our hunch that such effects, if they existed, would likely be small in size, so we pooled data across observers to emphasize mean differences. For a given experiment, we then compared matched conditions across the two portions of data using randomization, to ascertain whether the degree of similarity in stimulus orientations had any obvious effects on performance.

2.2.5 Eye-movement recording

In each session, eye position (right eye, 500 Hz) was recorded using an Eyelink 1000 (SR Research) and analyzed offline using custom MATLAB routines. Before each block, a cali-

bration routine (5-pt or 9-pt) was run. Trial onset was controlled in a gaze-contingent manner, beginning only after fixation was maintained within 2° or 2.5° of the fixation cross for 250 ms (anecdotal evidence suggested that drift was not uncommon towards the end of a block, hence we used a relatively large cut-off radius to limit unnecessary trial disruption). As our experiments involved relatively long duration trials and blocks, most observers systematically blinked during response and inter-trial intervals to limit eye fatigue. To focus our analysis on intervals of interest, we analyzed only position data from the onset of the first stimulus pair until disappearance of the second pair (1.4 s total). The trial-by-trial saccade detection proceeded as follows, closely following default Eyelink criteria and other well-accepted conventions (Engbert & Kliegl, 2003): from trial onset to first stimulus onset, we calculated the median horizontal and vertical eye position; these values were subtracted from the position data within the analysis window, so as to limit the effect of recording drift across individual blocks. Velocity along the horizontal and vertical axes was calculated by applying a sliding 5-pt window to the position data, and Euclidean velocity was then calculated. Euclidean acceleration was calculated in a similar fashion. To avoid contaminating the saccade detection analysis with blinks, samples that corresponded to blinks (and 100 ms either side) were removed, by searching for intervals where pupil size data was not recorded. Saccades were detected by searching for samples where velocity exceeded 30°/s, peak acceleration exceeded 8000°/s^2, and amplitude (i.e., Euclidean distance from rising above to falling below 30°/s) exceeded 0.5°. We repeated the randomization analyses of the main experiments after excluding trials containing saccades during the analysis window (the percentage of which ranged from approximately 1% to 12.5% of total trials in different observers, using this relatively small saccadic cut-off). Eye data from a small number of blocks

(four out of several hundred) was accidentally overwritten during the course of running experiments.

2.2.6 Control experiment

In a control experiment, we replicated the main contrast-discrimination experiment, this time providing separate response button pairs for each hand (four-button task). Data were collected from an overlapping group of eight observers (two authors), with one additional recruit withdrawing after completing the practice and first test session of the task. Participants were instructed to discriminate in which interval the target grating had higher contrast, this time using the response keys on the side in which the target appeared (indicated by the post-cue) e.g., pressing the nearer of the two buttons ('v' or 'n') for interval 1, or the further of the two buttons ('f' or 'j') for interval 2. This allowed us to estimate thresholds using only trials on which the observer explicitly indicated having responded towards perceived changes at the target location (note that non-target responses were considered incorrect responses, and still modified the staircase position). This procedure also allowed us to estimate the relative frequency of non-target responses across different conditions.

2.3 Results

2.3.1 Sensory discrimination in the presence of distractors

Performance across experiment differed in terms of its distractor-dependence, as illustrated by the mean thresholds across observer (Figure 2.2B and C). Mean contrast-discrimination thresholds were estimated to be around 10% contrast for the low and intermediate (10, 20, and 40%) distractor contrast conditions, but were approximately doubled for targets paired

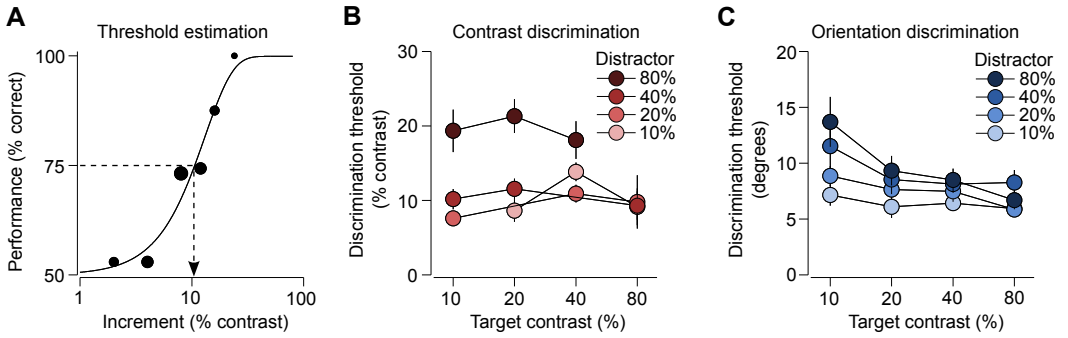

Figure 2.2: Sensory discrimination in the presence of distractors: data. A) Example psychometric function, and depiction of threshold estimation procedure. B) Mean contrast- and C) orientation-discrimination thresholds across observers (n = 8). Individual curves represent threshold as a function of target contrast, plotted separately for each distractor contrast level. The same eight observers completed both experiments. Bars represent standard error across observers.

with the highest contrast distractor (Figure 2.2B). This increase in threshold was not restricted to the lowest contrast target (10%), also occurring for targets of 20% and 40% contrast. Note that we avoided target contrast values in the very low range, where monotonic increase of threshold is typically most evident, and that the relatively high baseline thresholds observed in this divided attention task are not at odds with previous findings (Pestilli et al., 2011). For orientation discrimination, thresholds were lower overall with higher target contrast (Figure 2.2C), albeit with higher baseline threshold than one would obtain under conditions without target location uncertainty (Skottun et al., 1987). Distractor contrast had comparatively weaker effects on orientation discrimination: while mean thresholds increased with each increase in distractor contrast, these effects were more graded in fashion, with no indication of a large threshold jump from 40% to 80% contrast distractor conditions.

These general conclusions were backed up by statistical test. At each target contrast level within each experiment, we performed pair-wise randomization tests via bootstrap-

ping, to estimate the probabilities of observing distractor-mediated differences in threshold by chance (Table 2.1). This pair-wise analysis illustrated comparatively greater distractor influence in the contrast experiment. For 10% contrast targets, for example, contrast-discrimination thresholds were much more affected by 80% compared to 40% ($p = 0.0001$) or 20% contrast distractors ($p = 0.0001$), while the probabilities of observing such differences for orientation-discrimination thresholds were comparatively weaker (80% vs 40%: $p = 0.1206$; 80% vs 20%: $p = 0.0016$; 80% vs 10%: $p = 0.0001$). Similarly for 20% contrast targets, contrast-discrimination thresholds were again much more strongly affected by 80% compared to 40% ($p = 0.0001$) or 10% contrast distractors ($p = 0.0001$), while the probabilities of observing such differences for orientation-discrimination thresholds were comparatively weaker (80% vs 40%: $p = 0.2373$; 80% vs 20%: $p = 0.0471$; 80% vs 10%: $p = 0.0024$).

To rule out a variety of extraneous factors as possible explanations of the data, we ran several control randomization analyses. First, for both datasets we re-calculated thresholds after removing trials containing saccades of 0.5° amplitude or greater (see Materials and Methods and Figure 2.3A). These accounted for approximately 1% to 12.5% of total trials in different observers, using this relatively small saccadic cut-off. The overall pattern of results across experiment was unchanged (e.g., for 10% contrast targets, probabilities of difference for 80% vs 40% distractors were: $p = 0.0001$ for contrast discrimination and $p = 0.3153$ for orientation discrimination; for 20% contrast targets, the probabilities were: $p = 0.0001$ and $p = 0.2545$ respectively). Thus, eye movements during the stimulus intervals are unlikely to explain the differences observed across experiments. In addition, we also re-calculated the main randomization p-values after excluding data whose psychometric function fit failed to

	Distractor pair					
	10% vs 20%	10% vs 40%	10% vs 80%	20% vs 40%	20% vs 80%	40% vs 80%
			Contrast discrimination			
Target						
10%	-	-	-	.0106	.0001	.0001
20%	-	.0288	.0001	-	-	.0001
40%	.9712	-	.0039	-	.0001	-
80%	.4060	.4654	-	.6264	-	-
			Orientation discrimination			
Target						
10%	.0771	.0031	.0001	.0584	.0016	.1206
20%	.0390	.0129	.0024	.1954	.0471	.2373
40%	.1243	.0283	.0147	.2616	.1709	.3595
80%	.5446	.0071	.1449	.0061	.1290	.9656

Table 2.1: The table presents randomization analyses on threshold differences as a function of distractor pair. P-values for each condition were read off from bootstrapped distributions with 10,000 samples each (see Materials and Methods). Smaller p-values indicate larger thresholds for the higher contrast distractor condition.

pass a two-tailed deviance analysis measuring goodness-of-fit (see Materials and Methods). The overall pattern of results was again similar (e.g., for 10% contrast targets, probabilities of difference for 80% vs 40% distractors were: $p = 0.0024$ for contrast discrimination and $p = 0.0263$ for orientation discrimination; for 20% contrast targets, the probabilities were: $p = 0.0001$ and $p = 0.1972$ respectively). Thus, factors such as observer fatigue, across-session learning, or experimenter bias in hand-picking increment values are unlikely to have played any causal role in the observed differences across experiment.

Finally, we also carried out two additional analyses on data pooled across observers (see Materials and Methods). First, we verified that the increased exposure observers had on the orientation-discrimination task (due to a pre-experimental training protocol) played

no obvious role in our observed experimental differences, by comparing subsets of the data for which observers had as much or more exposure to the contrast-discrimination task relative to the orientation-discrimination task. Effect sizes were comparable to results of our main analyses (e.g., for 10% contrast targets, probabilities of difference for 80% vs 40% distractors were: $p = 0.0001$ for contrast discrimination and $p = 0.0582$ for orientation discrimination; for 20% contrast targets, the probabilities were: $p = 0.0002$ and $p = 0.6315$ respectively). Second, we verified whether distractor orientation played any role in the measured threshold behaviors. To do this, we split the datasets into two parts, trials in which target and distractor had closer to parallel orientations, and trials in which stimuli were closer to orthogonal. We then re-calculated thresholds for each set of trials separately, and computed the relevant threshold differences (thresholds for orthogonal trials minus thresholds for parallel trials, computed separately for each target-distractor condition). The resulting threshold differences were generally small for both tasks, as reflected in a randomization analysis; yet, there did appear to be a consistent orientation similarity effect in the orientation-discrimination task, with thresholds for closer to parallel trials smaller on average than thresholds for the orthogonal pair trials, and with most benefit at lower contrast targets (Figure 2.3B).

2.3.2 Control experiment

Results from the main experiments support the notion of sub-optimal response pooling in contrast-discrimination tasks with target location uncertainty, in agreement with at least one prior finding (Pestilli et al., 2011). Yet, our experiments cannot rule out the possibility that observers may have responded on some sizeable fraction of trials to perceived contrast

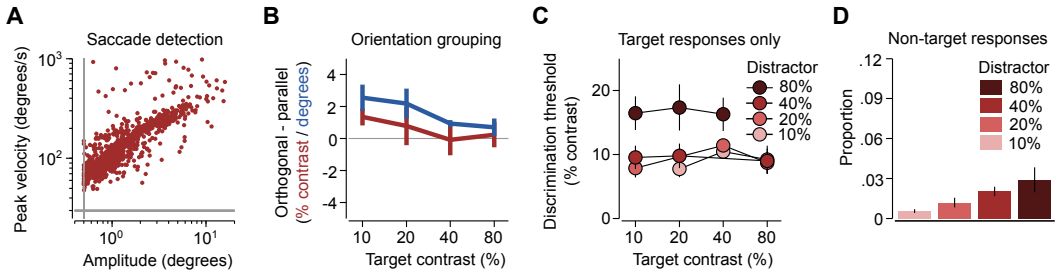

Figure 2.3: Sensory discrimination in the presence of distractors: control analyses. A) Main sequence data from the contrast-discrimination experiment (pooled data). Thresholds were re-computed after removing saccade trials (see main text for details of randomization analyses). B) Threshold difference between trials with orthogonal vs. parallel orientation pairs. Thresholds were calculated from pooled observer data (see Materials and Methods). Positive values indicate larger thresholds for trials in which the orientation pairs were closer to orthogonal. At each target contrast, mean and s.e.m. are calculated across the three (contrast) or four (orientation) levels of distractor contrast. C) We replicated the main contrast-discrimination experiment, this time providing observers with two response buttons for each hand, and instructing them to respond with the target-sided hand only. Mean thresholds (n = 8) from target response trials only are illustrated. D) Proportion of non-target responses as a function of distractor contrast, averaged over target contrasts and s.e.m. taken across observers.

changes at the distractor location; alternatively, observers may have adopted some explicit strategy whereby they attempted to represent the average contrast of both items in a given interval, and then respond to the interval with the higher average contrast. In an attempt to better understand the strategy employed by observers in the contrast-discrimination task, we replicated our experimental design with a new group of observers, this time providing observers with separate response button pairs for each hand, and instructing them to respond using the target-sided hand only. This allowed us to estimate thresholds using only trials on which the observer explicitly indicated having responded towards the target location. In addition, it allowed us to estimate the relative frequency of non-target responses across the different conditions.

Data are illustrated in Figure 2.3C and D. Observers performed the task successfully, responding using the target-sided response keys for the vast majority of trials. Thresholds estimated from the target-sided responses were highly distractor-dependent, replicating the

general shape of the data from the main contrast-discrimination experiment (Figure 2.3C). A randomization analysis confirmed that effect sizes were comparable to the earlier results (e.g., for 10% and 20% contrast targets, the probability of difference for 80% vs 40% distractors was in both cases: $p = 0.0001$). Nevertheless, there did appear to be a weak tendency on average for more non-target responses in higher-contrast distractor conditions, suggesting some small amount of above-threshold driving of responses by the distractor (Figure 2.3D). Overall, however, results of the control experiment suggested that responses driven by the non-target stimulus location likely played no substantial role in the high-contrast distractor effect. It is worth pointing out, however, that this control experiment might have undersampled the true proportion of trials on which decision was influenced by perceived changes at the distractor location. That is, the instructions explicitly emphasized for observers to give target-sided responses; thus, occasionally observers may have given target-sided responses despite decision being driven by perceived changes at the distractor location. Such target-distractor interactions, however, would presumably be very difficult to disentangle, either experimentally or conceptually, from the notion of sensory pooling.

2.4 Conclusion

We investigated how observers select sensory information in performing contrast and orientation discriminations, by measuring the effects of high-contrast distractors (i.e., large sensory responses) on behavioral performance in these tasks. Prior work had shown that when selecting from multiple stimuli that vary in contrast, distractors that evoke large sensory responses severely impact contrast-discrimination performance (i.e., lead to larger thresholds), supporting a model of sensory selection in which sensory responses are sub-optimally

pooled across space (Pestilli et al., 2011; Chen & Seidemann, 2012; Hara & Gardner, 2014). We observed large increases of contrast-discrimination thresholds when targets appeared in the presence of a single high-contrast distractor placed in the opposite hemifield. For orientation discrimination, however, high-contrast distractors had relatively moderate effects on performance at a target location, disrupting thresholds in a weaker, graded fashion.

What factors underlie the seemingly incommensurate behavior we observed in the separate experiments? In either task, an ideal observer would retain an estimate of the relevant stimulus property from both stimulus locations during interval one, repeat this process for interval two, and compare the difference in estimates across interval. Yet, fundamental differences in how estimates of contrast and orientation are encoded and maintained over time likely give rise to the profiles of threshold behavior we observed. For example, contrast-discrimination performance with single apertures is known to fall off rapidly with response delay periods of only a few seconds, while orientation discrimination is little affected for delays of 10s or more (Lee & Harris, 1996; Magnussen & Greenlee, 1999). Having to retain estimates of multiple contrasts simultaneously, as in the present task, would presumably burden visual short-term memory to a greater extent, perhaps leading to noisier estimates or poorer separation of individual estimates in memory.

In fact, when we consider the present results alongside several other recent findings (Pestilli et al., 2011; Sergent et al., 2011; Xing et al., 2014), it is difficult to escape the following conclusion: observers appear ill-equipped to represent and store more than a single contrast estimate at a time (Pestilli et al., 2011; Xing et al., 2014); yet, multiple orientation estimates at a time can be easily stored and inspected from memory (Sergent et al., 2011). Such striking differences in short-term memory performance are presumably underpinned

by more fundamental encoding or VSTM differences for the two types of sensory information (Xing et al., 2014). In the following chapter, we will fit a neural population model to the data from these experiments, in an effort to better understand the likely source of the differences.

In conclusion, we investigated the task-dependent nature of sensory selection, by testing the effects of large sensory responses on observer performance in two standard visual-discrimination tasks. It is well accepted that individual neurons involved in decision-making likely receive inputs from sensory neurons with widespread retinotopic locations and feature selectivities. In line with recent experimental findings (Pestilli et al., 2011; Chen & Seidemann, 2012), we found evidence that selection during contrast discrimination is severely disrupted by the presence of large sensory responses elsewhere in the visual field. In judging orientation changes, however, observers appeared to encode and maintain the relevant information more precisely in visual short-term memory, with high-contrast distractors having comparatively weaker effects on decision.

3

Efficient and inefficient selection from the same sensory neural response: computational model

3.1 Introduction

Disentangling the roles of sensory, memory and decision-related factors on psychophysical performance is an inherently convoluted exercise, requiring tightly controlled experiments alongside equally well-formulated models. Key insights on the nature of stimulus encoding and decoding have been gained, for example, by fitting precise neural com-

putational models to data from sensory discrimination tasks (Chirimuuta & Tolhurst, 2005; Seriès et al., 2009; May & Solomon, 2015). As a metric of discrimination behavior, neural population model approaches have often utilized the concept of Fisher information (FI), a measure of the best possible decoding performance obtainable for an unbiased estimator of an encoded stimulus value (Paradiso, 1988; Seung & Sompolinsky, 1993; Dayan & Abbott, 2001). By providing a precise bound on the decoding accuracy possible for a given encoding architecture, FI-based models have shed light on key factors limiting decoding performance in sensory discrimination tasks, such as neural adaptation (Seriès et al., 2009), noise correlations (Averbeck & Lee, 2006; Ecker et al., 2011) and stimulus priors (Ganguli & Simoncelli, 2014).

The computations underpinning sensory discrimination for single, isolated stimuli are now relatively well explored, with the existence of a number of well-developed decoding models for stimulus features such as contrast and orientation (Paradiso, 1988; Graf et al., 2011; Sanborn & Dayan, 2011; May & Solomon, 2015). Decoding performance in the context of target uncertainty, however, is still a very poorly understood problem. In one prominent approach, selection of the maximum of a set of neural responses acts as a coarse proxy for decision, so called max-pooling models of sensory selection (Pelli, 1985; Palmer et al., 2000; Pestilli et al., 2011). While max-pooling models have been successful in fitting contrast-discrimination thresholds under conditions of target uncertainty (Pestilli et al., 2011; Hara & Gardner, 2014), such decision rules are somewhat ad-hoc in nature, and are known to perform poorly relative to ideal observer models for tasks such as visual search and change detection (Ma et al., 2015). In addition, it is not always clear how max-pooling models overlap with other models of sensory interaction and attentional selection, such as

models based on divisive normalization (Heeger, 1992; Reynolds & Heeger, 2009; Itthipuripat et al., 2014). In sum, there is a need for computational approaches that can more readily explore the interactions between multiple stimuli across potentially multiple feature dimensions.

Some recent efforts have been made in this regard; yet, these have either been purely theoretical in nature (Orhan & Ma, 2015), or have been focused on the stimulus estimation paradigm (Matthey et al., 2015). One potentially promising line of investigation relates to the notion of linear mixing of neural responses. Numerous investigations have now illustrated how neural responses to a given stimulus are often well-approximated by a weighted sum of neural responses to the individual, constituent features of that stimulus (Recanzone et al., 1997; Zoccolan et al., 2007; Busse et al., 2009). A recent theoretical investigation suggests that this type of neural response mixing may be substantially more detrimental to decoding performance than sizeable decreases in response gain or noise amplitude (Orhan & Ma, 2015). Thus, a linear mixing framework may be particularly appropriate for understanding the strength of distractor effects in different stimulus contexts and tasks, and given the right formal decoding approach (i.e., FI), might be readily studied across multiple decision spaces simultaneously.

In Chapter 2, we systematically tested the effects of distractor contrast on the discrimination of changes to the contrast or orientation of a target stimulus. Contrast-discrimination performance was severely disrupted when high-contrast distractors appeared in the opposite visual hemifield to the target, while disruption of orientation-discrimination performance was more graded in magnitude. In the present chapter, we aimed to develop a computational understanding of the nature of these stimulus interactions. To do so, we con-

structed an FI-based model for the simultaneous estimation of contrast- and orientation-discrimination thresholds from an idealized neural population. We validated the form of our encoding model and the relevance of the FI approach by first fitting the model to single-stimulus discrimination data collected in a separate auxiliary experiment. An encoding model with multiple neural subpopulations and an expansive population-level contrast response successfully replicated the single-stimulus discrimination data. We then extended this model to allow for interactions between target and distractor stimuli; for the particular tuning parameterizations we explored, however, neither of two common models of sensory interaction (divisive normalization, linear mixing of neural responses) could satisfactorily explain the datasets from Chapter 2 simultaneously. We discuss possible reasons for the failure of these models, in the process suggesting a number of possible future extensions of this work.

3.2 Materials and Methods

3.2.1 Background

In Chapter 2, we measured contrast- and orientation-discrimination thresholds for target stimuli presented alongside an irrelevant distractor. The cornerstone of our modeling efforts in this chapter is the computation of FI in an idealized sensory neural population that simultaneously encodes stimulus contrast and orientation. FI is a measure of the maximum accuracy achievable by an unbiased estimator of an encoded stimulus value, and is equal to the inverse of the square of the discrimination threshold (Seung & Sompolinsky, 1993; Abbott & Dayan, 1999).

Our approach consisted of two stages. First, we validated the FI approach by fitting a

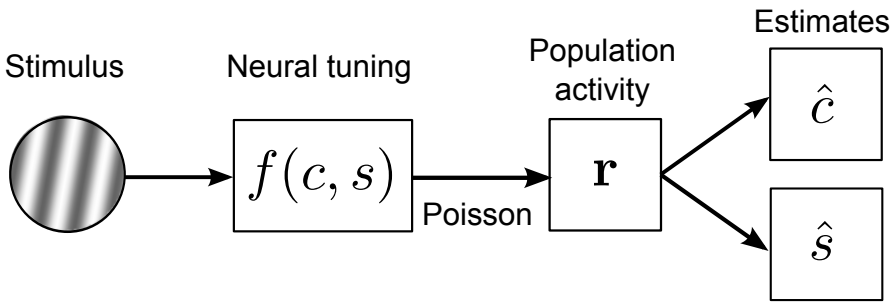

Figure 3.1: Sensory discrimination for isolated stimuli: encoding-decoding model. To validate the Fisher information approach, we developed a neural population model for predicting contrast- and orientation-discrimination thresholds for single, isolated stimuli. See Materials and Methods for details.

model to single-stimulus contrast- and orientation-discrimination thresholds collected in a separate, auxiliary experiment. This allowed us to understand the basic encoding model architecture necessary for fitting these types of discrimination data simultaneously. The model architecture we define at this stage resembles one other recent approach (May & Solomon, 2015). We then adapted this conceptual framework with the goal of accounting for the distractor-dependent threshold increases described in Chapter 2. To do so, we allowed for interactions between spatially separated target and distractor stimuli, comparing the performance of two very general models of sensory interaction, divisive normalization (Heeger, 1992; Carandini & Heeger, 2012) and linear mixing of neural responses (Orhan & Ma, 2015).

3.2.2 Neural model: Discrimination thresholds for isolated stimuli

Fisher information

We consider the problem of simultaneously encoding contrast c and orientation s in a population of independent Poisson-like neurons (Figure 3.1). We assume that the mean re-

sponse of the i-th neuron to (c, s) is given by:

$$f_i(c, s) = g_i(c)h_i(s), \tag{3.1}$$

where $g_i(c)$ controls the contrast response of the neuron and $h_i(s)$ determines the neural tuning to stimulus orientation. For the contrast response function, $g_i(c)$, we select a monotonic equation of Naka-Rushton form (Naka & Rushton, 1966),

$$g_i(c) = \frac{c^{n_i}}{c^{n_i} + \alpha_i^{n_i}}. \tag{3.2}$$

The responsiveness of the neuron to contrast is governed by the exponent, n_i, and the semi-saturation contrast, α_i. For orientation tuning curves, we select a population of homogeneous Von Mises functions,

$$h_i(s) = \beta_i \exp(\gamma_i(\cos(s - s_i) - 1)) \tag{3.3}$$

where the parameters β_i and γ_i determine the response gain and the concentration parameter (narrowness) of the tuning curve. We assume independent Poisson noise:

$$p(\mathrm{r}|c, s) = \prod_i p(r_i|c, s) \tag{3.4}$$

$$= \prod_i \frac{1}{r_i!} e^{-f_i(c,s)} f_i(c, s)^{r_i}. \tag{3.5}$$

To compute model thresholds, we first need to compute the Fisher information matrix (FIM):

$$\mathrm{I}(c,s) = \begin{pmatrix} I_{cc} & I_{cs} \\ I_{cs} & I_{ss} \end{pmatrix}. \tag{3.6}$$

This 2x2 matrix has three components: I_{cc} and I_{ss} on the diagonal, and I_{cs} off-diagonal. The components of the FIM are computed in the standard way (Dayan & Abbott, 2001). For example,

$$\begin{aligned} I_{cc} &= -\left\langle \frac{\partial^2 \log p(\mathrm{r}|c,s)}{\partial c^2} \right\rangle & (3.7)\\ &= -\sum_i \left\langle \frac{\partial^2 \log p(r_i|c,s)}{\partial c^2} \right\rangle, & (3.8) \end{aligned}$$

where $\langle\cdot\rangle$ is the expected value under $p(\mathrm{r}|c,s)$. This expression can be evaluated as

$$I_{cc} = \sum_i \frac{\left(\frac{\partial f_i(c,s)}{\partial c}\right)^2}{f_i(c,s)}. \tag{3.9}$$

Similarly,

$$\begin{aligned} I_{ss} &= \sum_i \frac{\left(\frac{\partial f_i(c,s)}{\partial s}\right)^2}{f_i(c,s)} & (3.10)\\ I_{cs} &= \sum_i \frac{\frac{\partial f_i(c,s)}{\partial c}\frac{\partial f_i(c,s)}{\partial s}}{f_i(c,s)} & (3.11) \end{aligned}$$

Per the Cramér-Rao bound (Cover & Thomas, 1991), the covariance matrix of an optimal estimator is the inverse of the FIM:

$$\Sigma(c, s) = \mathrm{I}(c, s)^{-1} \tag{3.12}$$

The predicted thresholds for contrast and orientation can then be computed as the square roots of the diagonal elements of the resulting covariance matrix.

Tuning properties

So far, we have computed optimal thresholds for an idealized neural population with defined contrast and orientation tuning. However, this tuning has free parameters. We now specify how we chose these parameters to tile the respective feature dimensions, adapting an approach recently taken by May & Solomon (2015) to fit discrimination thresholds.

First, we took a population of N = 256 neurons, and divided this into 8 smaller subpopulations of N = 32 neurons each. Within each subpopulation, the preferred orientations of the neurons were set to fixed values equally spaced in orientation space. The parameter controlling the narrowness of the orientation tuning width, γ, was allowed to vary in model search, but was held fixed across all neurons in the entire population. The remaining parameters were held fixed within each subpopulation, but were allowed to vary across the subpopulations within certain constraints. Semi-saturation contrasts, α, were spaced in equal logarithmic steps between 0.01 and 3 (proportion contrast); values greater than 1 have been observed in neural recordings (Albrecht & Hamilton, 1982) and have been used in prior model efforts similar to ours (Chirimuuta & Tolhurst, 2005; May & Solomon, 2015). For the remaining parameters, we explored two scenarios:

- β and n were held constant for all subpopulations (constant parameterization).
- β and n were allowed to vary by subpopulation index j, according to independent, two-parameter exponential functions (expansive parameterization).

These choices were motivated by a couple of desires. First, having a potentially graded scaling of maximum firing rate for each subpopulation seemed logical, as in the normalized case, the idealized contrast tuning functions naturally asymptote at different response amplitudes. Similar approaches have been used previously (Chirimuuta & Tolhurst, 2005; May & Solomon, 2015). Second, as has now been documented a number of times, model fits to contrast-discrimination thresholds seem to require an expansive, population-level contrast response, to account for the flattening or dip in thresholds at very high-contrasts (Kingdom & Whittle, 1996; Zenger-Landolt & Heeger, 2003; Chirimuuta & Tolhurst, 2005; May & Solomon, 2015). From a modeling perspective, this could be incorporated in a number of ways; we chose an approach where the relevant contrast sensitivity parameters were either held fixed, or were allowed to increase in magnitude with the semi-saturation contrast of the subpopulation. In model fits, the constant and expansive model variants above had 3 free paramters (β, n, γ) and 5 free parameters (γ, plus two independent, 2-parameter exponential functions controlling β and n) each.

3.2.3 Neural model: distractor effects on sensory discrimination

Models of sensory interaction

We then extended our FI model approach to incorporate either of two general models of sensory interaction, in an attempt to fit the data from Chapter 2. In one version of the

model, we allowed for sensory interactions through the linear mixing of separate, sensory neural responses (Orhan & Ma, 2015). In this model, we assume two neural populations encoding the contrasts and orientations of the target and distractor gratings. For neurons in the first group, the mean response is given by:

$$f_i(c_t, c_d, s_t, s_d) = w_1 \frac{c_t^{n_i}}{c_t^{n_i} + \alpha_i^{n_i}} h_i(s_t) + w_2 \frac{c_d^{n_i}}{c_d^{n_i} + \alpha_i^{n_i}} h_i(s_d) \tag{3.13}$$

where c_t and c_d are the target and distractor contrasts, and s_t and s_d are the target and distractor orientations. We assume $w_1 > w_2$ for neurons in this group, hence they primarily encode the target grating; for neurons in the other group, the weights are switched, so they primarily encode the distractor. We further assume $w_1 + w_2 = 1$. We can think of w_1 and w_2 as roughly capturing receptive field effects, for example, in some higher-level read-out stage where spatial receptive fields are broad.

We also tested a model in which sensory interactions occurred through divisive normalization (Heeger, 1992; Carandini & Heeger, 2012). For the divisive normalization model, we used an architecture like that described above for the linear mixing model, but with two modifications. First, we set the weight term in the model to 1, so that there was no response mixing (although there was still global read-out from separate target-centered and distractor-centered neural populations). Second, we incorporated the relevant normalizing or divisive terms into the model i.e., for the target-centered population, the distractor contrast was added to the denominator of the Naka-Rushton gain expression; for the distractor-centered population, the target contrast was added to the denominator. In model fitting, we allowed these divisive terms to be scaled by a flexible weight term, k.

We selected these models for a number of reasons. Divisive normalization, for exam-

ple, is thought to be a canonical computation in sensory and neural systems (Heeger, 1992; Carandini & Heeger, 2012), and may reflect the broader feedforward inhibition typically found in neural circuits. In the present context, it could arise through rapid fluctuations in attentional gain. Linear mixing, on the other hand, has been observed in numerous visual computations (Recanzone et al., 1997; Zoccolan et al., 2007; Busse et al., 2009), and is an appropriate choice of model where graded stimulus interactions are concerned (Orhan & Ma, 2015). Linear mixing might also capture behaviors similar to max-pooling (i.e., for high-contrast distractor conditions).

As with the earlier model, we assumed independent, Poisson-like noise. We simulated model performance by computing all sixteen entries of the 4x4 FIM according to:

$$I_{xy}(c_t, c_d, s_t, s_d) = \sum_i \frac{\frac{\partial f_i}{\partial x}\frac{\partial f_i}{\partial y}}{f_i(c_t, c_d, s_t, s_d)} \tag{3.14}$$

where the x and y pair takes on each possible combination of the indices: (c_t, c_d, s_t, s_d). After computing these terms, we then estimated thresholds by inverting the FIM, and taking the square root of the resulting diagonal variance terms as before. In passing, note that the sum across neurons described here and earlier is the same whether or not the neurons are also indexed by subpopulation.

Tuning properties

For both sensory interaction models, we adopted the same expansive model architecture used in fitting the single-stimulus discrimination data. Specifically, we allowed for N = 256 neurons spread across 8 subpopulations. Semi-saturation contrasts were again set to fixed

values that varied from subpopulation to subpopulation, tiling the contrast axis in equal logarithmic steps between fixed bounds of 0.01 and 3 (proportion contrast). Each of the sensory interaction models had a total of six free parameters (four exponential function parameters, γ and, either w or k).

3.2.4 Model fitting

In both phases, we fit models to contrast- and orientation-discrimination thresholds simultaneously (i.e., constrained by the measured thresholds from both experiments), searching for the set of parameters that gave the smallest combined RMSE relative to observed thresholds. To calculate combined RMSE, observed thresholds were converted to proportions (0-1) and radians respectively, and the squared error relative to model predictions was calculated for each test condition before taking the mean and square root. For the single-stimulus discrimination data, this included ten conditions. For the two-stimulus data from Chapter 2, this included twenty-eight conditions in total (twelve contrast and sixteen orientation conditions). In all model fits, the preferred orientations of the neurons within a given subpopulation were equally spaced in orientation space between $-\pi$:π, and for simplicity, we assumed the orientations of the target and distractor stimuli were centered in this space.

We also made two minor transformations of the predicted thresholds prior to RMSE calculation. First, the thresholds measured in Chapter 2 were estimated from the 75% correct performance point taken from Weibull model fits; however, the σ value returned by FI reflects a different performance level (i.e., approximately 84.1% correct performance on a cumulative gaussian). Concerned that this deviation might affect our optimizations in some

systematic way, we approximated the 75% percent correct point prior to RMSE calculation, by rescaling the predicted thresholds appropriately. In addition, as the model orientation space ($-\pi$:π) was twice the size of the space in which orientation-discrimination thresholds were actually measured, we halved the predicted orientation-discrimination thresholds before RMSE calculation.

In all model-fitting, we used an evolutionary search algorithm known as Covariance Matrix Adaptation (CMA-ES) (Hansen & Ostermeier, 1996). This is a very robust, cutting-edge optimization algorithm that makes few assumptions regarding the nature of the function being optimized. We implemented CMA-ES using a freely available MATLAB function version (*cmaes.m*, available at www.lri.fr/∼hansen/cmaesintro). As a stochastic algorithm, CMA-ES gives different output depending on the random seed used for model initiation. For all model fits, we ran 100 separate searches using the high-performance computing (HPC) cluster at New York University, each with a different set of starting parameter values for the optimization. Best-fit parameters for a particular observer/model were taken from the model run that gave the smallest RMSE out of all runs. Individual parameters were allowed to vary within broad but finite bounds as follows: the scale of the two exponential gradients (or constant in the case of the constant parameterization fits to single-stimulus data) were: β: 1-100 and n: 0.1-4. The power of each exponential gradient could vary from 0:5, thus allowing for potentially expansive increase in parameter value with increasing semi-saturation contrast. Remaining parameters could range as follows: γ: 0.1-6; w (linear mixing): 0.5-1; k (divisive normalization): 0-1. We added a small error term to the lower bound for w, to prevent the emergence of singular and badly-scaled matrices in the optimization. Note also that the narrowest possible tuning function width ($\gamma = 6$) is

equivalent to a gaussian σ of approximately 23.4°, reasonable given that 32 neurons tiled the model orientation space of $-\pi$:π. For all model fits, optimization starting values were randomly drawn from within the 25th-75th percentile ranges for each of these parameter ranges, and the CMA-ES search width parameter was set to one-third of the respective parameter range.

3.2.5 Auxiliary experiment

Data collection for the main experiments is described in detail in Chapter 2. Below, a brief description is given for an auxiliary experiment measuring single-stimulus discrimination thresholds. This was carried out with the aim of appropriately constraining computational models of performance in the main experiments, and is thus described here. The general methods and informed consent procedures were largely identical to those described in Chapter 2.

In the auxiliary experiment, we measured contrast- and orientation-discrimination thresholds for isolated targets, using a largely identical set-up to the main experiments (with distractor contrast now set to 0% contrast). Data were collected from seven observers (in two or three separate sessions each), including three (two authors) who completed the main experiments. Distributed cues (i.e., white arrows) were presented before and during the stimulus intervals, such that observers did not know in advance on which side the stimulus would appear (although the target did appear on the same side for both intervals, so we can assume the second interval location was known). A post-cue again indicated target location. As a primary goal of this experiment was to collect data that might validate our model approach, we included one additional low-contrast target condition (2%), so as to better mea-

sure the dynamic range of discrimination thresholds. Attempts to measure performance at much lower pedestal contrast were abandoned, as we felt that orientation discrimination at much lower contrast and with spatial uncertainty amounted to a form of stimulus detection. Thresholds were estimated using procedures similar to those described earlier.

3.3 Results

3.3.1 Fisher information-based model of discrimination thresholds

An important aspect of our experimental design in Chapter 2 was the use of identical stimulus parameters across experiment, varying only the task performed by the observer. By assuming that identical sensory neural responses were evoked across experiment, characteristics of how the individual sensory neural responses interacted and were decoded must underlie the threshold differences we observed. To formalize this approach to fitting the data, we developed an encoding-decoding model based on Fisher information (FI), for the simultaneous estimation of contrast- and orientation-discrimination thresholds from an idealized sensory neural response (Figure 3.1). In doing so, we make the simplifying assumption that thresholds are inversely related to the precision of the decoder, and exploit the fact that FI provides a measure of this precision. Specifically, subject to the Cramér-Rao bound, FI sets a lower limit on the accuracy with which the true stimulus value can be decoded by any unbiased estimator (Cover & Thomas, 1991; Abbott & Dayan, 1999). We validated the form of this model by first fitting it to single-stimulus contrast- and orientation-discrimination thresholds collected in a separate auxiliary experiment. Model fits were doubly constrained by fitting simultaneously to both contrast- and orientation-discrimination datasets (i.e., the model parameters were held fixed across tasks).

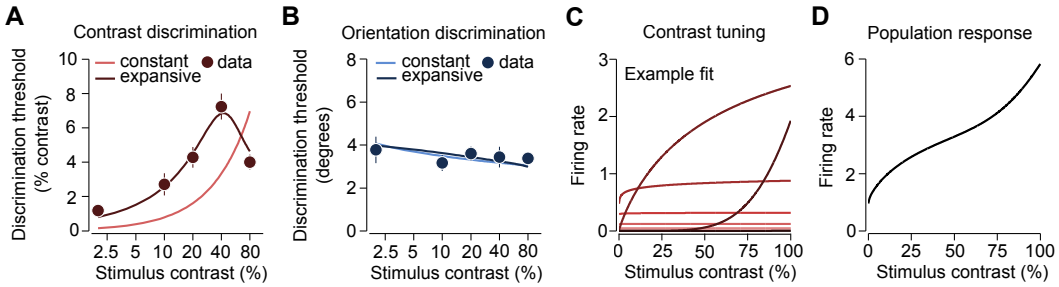

Figure 3.2: Sensory discrimination for isolated stimuli: model fits. A) Contrast- and B) orientation-discrimination thresholds were collected in an auxiliary experiment. Mean (s.e.m.) of the observed thresholds (n = 7) are plotted by the datapoints. Curves represent the mean of the predicted model thresholds from fits to the individual observer data. C) Predicted contrast tuning functions for the expansive parameterization model, for one illustrative observer. Darker colors represent subpopulations with higher semi-saturation contrast. For a given subpopulation, the shape of the function is controlled by the semi-saturation contrast, the contrast exponent, and the maximum firing rate. D) The summed population-level response illustrates a clearly expansive profile.

For isolated sensory neural responses, a model with contrast sensitivity parameters held fixed across neural subpopulations (constant parameterization) failed to satisfactorily fit real data (Figure 3.2A and B, RMSE mean/s.e.m. = 0.0212/0.0026). The single-stimulus contrast-discrimination thresholds exhibited a pointed decrease in magnitude for very high-contrast pedestals, a characteristic which has been documented before by others, under a variety of stimulation conditions (Kingdom & Whittle, 1996; Zenger-Landolt & Heeger, 2003; Pestilli et al., 2011). By designing model neurons such that they had logarithmically-spaced semi-saturation contrasts, while keeping all other parameters constant, model behavior was essentially Weber-like in form, with predicted thresholds increasing across the full range of the contrast axis. Thus, the model with constant parameterization for the contrast sensitivity parameters did not have sufficient flexibility to fit the late dip in thresholds. Presumably, the population-level contrast response relevant for the task is expansive at high-contrast values, which would result in decreased discrimination thresholds.

To build greater flexibility into the encoding model, we adapted an approach taken re-

cently by May & Solomon (2015). In this investigation, the authors accounted for the so-called 'near-miss' to Weber's Law by allowing for an exponential parameterization governing the contrast sensitivity of the population (May & Solomon, 2015). We incorporated this feature into our model, thus allowing for an expansiveness to the population-level contrast responses; preliminary model attempts suggested that flexibility of this sort was essential to fit the relatively sizeable late decrease in contrast-discrimination thresholds. It is worth pointing out, however, that a variety of methods might achieve the same goal. For example, appropriately designed transformations of the input luminance values (Chirimuuta & Tolhurst, 2005), or incorporation of a late expansiveness into the definition of the contrast response function itself (Zenger-Landolt & Heeger, 2003) could achieve qualitatively similar effects. In the model with expansive parameterization, we allowed the maximum firing rate and contrast exponent parameters (β, n) to increase as a function of semi-saturation contrast according to independent, 2-parameter exponential functions. In this way, we introduced a gradient of sensitivity across the subpopulations that would allow for an expansive population-level contrast response to emerge.

The model with expansive parameterization provided substantially better simultaneous fits to the single-stimulus threshold data, replicating the late threshold decrease for high-contrast stimuli (Figure 3.2A and B, RMSE mean/s.e.m. = 0.0083/0.0010). This behavior was driven by the substantially greater expansiveness in the contrast responses for the expansive model, as nicely depicted in an example of fitted tuning functions and summed population response for one illustrative observer (Figure 3.2C and D). This late expansiveness is qualitatively very similar to inferred contrast response functions found elsewhere e.g., in Fig. 6B of Zenger-Landolt & Heeger (2003).

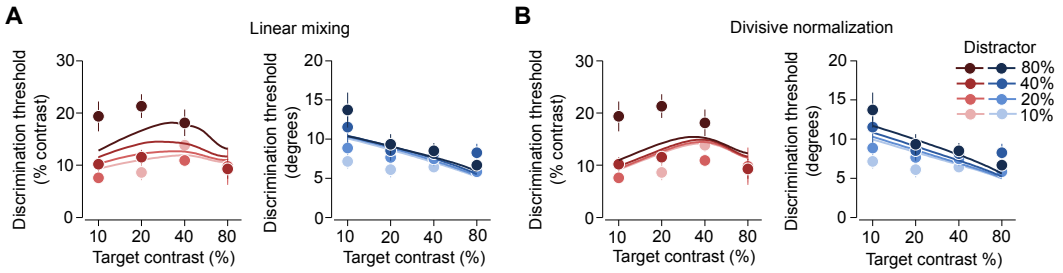

Figure 3.3: Sensory discrimination in the presence of distractors: model fits. A) Fits of the linear mixing model to the data from Chapter 2. B) Fits of the divisive normalization model to the data from Chapter 2. In each panel, curves represent the mean of the predicted thresholds from fits to the individual observer data. See Materials and Methods for details.

3.3.2 Sensory discrimination in the presence of distractors: model fits

The FI approach tested thus far is inherently local in application. Thus, conceptually at least, it may not be the appropriate level at which to implement models where target and distractor interact across broadly spaced locations and across two feature dimensions. In an attempt to fit the datasets illustrated in Chapter 2, we adopted an approach based on a recently described linear mixing model, in which FI is computed from a neural population that receives mixed inputs (Orhan & Ma, 2015). In this scheme, the readout neural population might be some higher-order sensory or decision-related neural population, where broader spatial receptive fields are likely. Such a scheme might allow for more graded and realistic distractor effects to emerge, while still qualitatively allowing for behavior not unlike more extreme models (i.e., max-pooling). In addition, we compared this model to a model in which distractor effects arose, instead, through divisive normalization (Heeger, 1992; Carandini & Heeger, 2012). For both models, we utilized the same tuning parameterization that was used for the single-stimulus model fits, thus allowing for a potentially expansive population-level contrast response.

Sensory discrimination in the presence of distractors: model fits

	S1	S2	S3	S4	S5	S6	S7	S8	Mean (s.e.m.)
Model									
LM	.0570	.0479	.0479	.0480	.0623	.0329	.0233	.0324	.0440 (.0047)
DN	.0516	.0586	.0471	.0554	.0637	.0358	.0240	.0361	.0465 (.0048)

Table 3.1: RMSE values for each observer from the best simultaneous fits to the data from Chapter 2. See Materials and Methods for details. LM = linear mixing model; DN = divisive normalization model.

Neither model provided satisfactory fits to the data (Figure 3.3). The linear mixing model, a weighted sum of independent (i.e., Naka-Rushton) sensory neural responses, failed to capture the shape of the two-stimulus data in general (Figure 3.3A). While the mean of the predicted contrast-discrimination thresholds for the linear mixing model did deviate as a function of distractor contrast, the magnitude of this effect was relatively small, and the model was unable to reproduce the observed orientation threshold increase with increasing distractor contrast. The model with divisive normalization also failed to replicate the shape of both datasets simultaneously (Figure 3.3B). In sum, both models provided quantitatively poor simultaneous fits to the two datasets, with neither satisfactorily recreating the profile of high-contrast distractor effects observed in Chapter 2.

When we look to the RMSE metrics from the individual observer fits (Table 3.1), we see little reason to favour either model definitively. At their simplest, these results suggest that model flexibilities other than purely linear mixing or divisive normalization are likely necessary to account for the shape of the two datasets simultaneously. We should also note that there appeared to be some degree of similarity in the performance of the two models, as assessed by the relatively similar RMSE values found for the models in a number of instances.

This may reflect some broad constraints that our tuning parameterization imposed on model behavior generally, perhaps preventing one or the other model from providing better fits. However, the same tuning parameterization very successfully fit the single-stimulus discrimination thresholds. While we lack detailed formal model validation and comparison in the analyses presented here, we did at least validate the tuning parameterization with a simpler dataset first, successfully fitting those data with an expansive tuning parameterization.

3.4 Discussion

Computational models of observer performance are essential components in any full understanding of task-specific psychophysical behavior. In the present chapter, we attempted to develop a precise neural population model of the psychophysical performance described in Chapter 2. We began by first validating an FI-based modeling approach, successfully fitting a specific encoding model simultaneously to contrast- and orientation-discrimination thresholds measured in a separate auxiliary experiment. We then extended this model to the more complex two-stimulus tasks described in Chapter 2, and found that neither of two general models of sensory interaction provided good simultaneous fits to those datasets.

In principle, FI-based models of sensory selection are an improvement on models based on sub-optimal, winner-take-all decoding. Nevertheless, key limitations of our approach must first be acknowledged. First, fits were not constrained by simultaneous neural response measurements, unlike in some prior related investigations (Pestilli et al., 2011; Ithipuripat et al., 2014). Second, the question of the physiological basis of the expansive contrast response is largely left unanswered here. While there is some evidence of broadly

divergent contrast sensitivities across magnocellular and parvocellular pathways (Kaplan & Shapley, 1986; Shapley, 1990), whether the late contrast threshold decrease we observed relates directly to some distinct neural source (i.e., the response of some subpopulation of neurons with high semi-saturation contrasts) is unknown. When optimized in fitting, the parameter values governing the sensitivity of the model to contrast varied substantially across the subpopulations i.e., with some individual values that could be considered unrealistically large or small when considered alongside analogous physiological measurements. Nonetheless, we found the reliance on expansiveness in the single-stimulus model fits broadly compelling, and suggest that the exact architecture of the tuning parameterization is of less importance, from the outset being abstract in nature.

The results here must also be examined for a variety of more technical reasons. First, we can not rule out the possibility that, with the addition of some simple, extra flexibility, one or the other sensory interaction models may have done a better job of fitting the two-stimulus datasets. For example, a combination of both forms of sensory interaction might be worth exploring in model form (i.e., a generalized linear mixing model with divisive normalization). In addition, we can not rule out the possibility that alternate forms of tuning function parameterization might have provided better grounding for model testing. For example, exploratory model searches led us to design an encoding model with a fixed, intermediate number of subpopulations (i.e., 8), and with the upper bound on the semi-saturation contrasts fixed above 1 (i.e., specifically, at the proportion contrast of 3). These were very specific design choices that on qualitative inspection and piloting, appeared to give relatively smooth model behavior relative to other discretizations of the space. As a result, however, we can not claim that the results here hold generally for all other architec-

tures. Yet, focus on such details could potentially overshadow the key point of the model fits to the single-stimulus discrimination data i.e., the necessity for the model to generate an expansive population-level contrast response. This type of response could be simulated in a variety of ways, with the abstract, underlying components utilized of less general interest. Future investigations should at least give some weight to the alternative approaches mentioned earlier e.g., models that transform the input luminance values (Kingdom & Whittle, 1996), or that add an expansiveness directly to the definition of the contrast response function (Zenger-Landolt & Heeger, 2003).

Finally, in deriving the FI-based expressions governing threshold behavior, we made several simplifying assumptions that should be mentioned. We chose subpopulations of set size and with homogeneous orientation tuning functions, and we assumed uniform noise correlation (set to zero) across the neural population. The possibility remains that some variant combination of gain model, tuning function and neural noise parameterization provides an alternative explanation of our datasets. For example, our assumptions about the form of neural noise may be limited in their validity, ignoring the large role now thought to be played by modulatory signals in setting the overall amplitude of sensory noise (Goris et al., 2014). While certainly over-simplistic, our assumption of zero correlation across neurons was a reasonable place to begin our model fitting. There is still much debate about how correlations modulate decoding performance, and about the conditions that determine whether increased correlations lead to facilitation or disruption of decoding (Abbott & Dayan, 1999; Ecker et al., 2011). Emphasis on such details here, however, would have obscured the main goal of testing two general models of interaction between separate sensory neural responses.

In conclusion, we studied the behavior of FI-based models of threshold performance, as simultaneously applied to data from the contrast- and orientation-discrimination experiments described in Chapter 2. We first validated this approach by fitting model expressions to single-stimulus discrimination data collected in auxiliary experiments. We then extended the model to allow for the linear mixing or divisive normalization of neural responses from separate target and distractor spatial locations. Neither sensory interaction model provided a quantitatively convincing fit to the datasets. Future model investigations might test alternate tuning function parameterizations, or combine both types of sensory interaction in a single model.

4

Delayed estimation of luminance contrast

4.1 Introduction

For many years, the contents of VSTM were conceptualized as being discrete in nature: an item was either in memory or it was not (Luck & Vogel, 1997; Cowan, 2001). From the point of view of perceptual psychophysics, this notion is simplistic to the point of being untenable. In signal detection theory models of psychophysical performance, for example, the observer's internal representation of a stimulus is taken to be a noise-corrupted version of the stimulus; thus, an item can be encoded to a greater or lesser degree, depending on the amount of noise in the representation. In the study of VSTM, however, it has

taken a long time for the concept of a memory being a noisy version of the stimulus to take hold. Cognitive psychology studies of VSTM have typically used coarse stimuli without parametric variation e.g., change detection among items that were hand-picked by the experimenter with the goal of making them highly discriminable (Pashler, 1988; Luck & Vogel, 1997; Cowan, 2001). In contrast, researchers working in the tradition of threshold psychophysics have long since adopted the idea of noisy memories, a concept implicit, for example, in paradigms that measure the magnitude of stimulus change necessary for some criterion level of discrimination performance (Palmer, 1990; Magnussen & Greenlee, 1999).

Recently, the concept of noisy memories has gained ground in VSTM research, due to the introduction of a new paradigm of probing VSTM known as delayed estimation (Wilken & Ma, 2004; Zhang & Luck, 2008; Fougnie et al., 2012; van den Berg et al., 2012; Bays, 2014). In this paradigm, inspired by earlier work from Prinzmetal and colleagues (Prinzmetal et al., 1997, 1998), the observer reports the identity of a remembered stimulus on a continuum, repeating this process over many trials to create a distribution of stimulus estimates. The width of the resulting estimate distribution can then be taken as a measure of the level of noise in the memory. Delayed estimation has been applied most notably to the study of VSTM for stimulus features such as orientation and color (Fougnie et al., 2012; van den Berg et al., 2012). For example, recent studies have quantified the dependence of noise level on set size, and attempted to determine whether there is an upper limit on the number of items that can be successfully remembered (Zhang & Luck, 2008; Bays et al., 2009; van den Berg et al., 2012).

While the vast majority of research using delayed estimation has studied features such as orientation and color, it cannot be assumed that these two features are representative of all

features stored in VSTM. For example, the neural representations of orientation and color rely on very specific neural substrates: topographically-arranged maps in the case of orientation (Hubel & Wiesel, 1962; Ferster, 2003), and specific color-opponent, retino-cortical pathways in the case of color (Lennie et al., 1990; Johnson et al., 2001; Gegenfurtner & Kiper, 2003; Brouwer & Heeger, 2009). If VSTM relies on the same neural networks responsible for initial sensory encoding, as some influential theories posit (Awh & Jonides, 2001; Jonides et al., 2008), then these structured neural representations provide an ideal substrate for relatively precise VSTM encoding and subsequent read-out. Thus, it is perhaps not surprising that accurate maintenance of orientation and color information is possible over relatively long delays (Nilsson & Nelson, 1981; Magnussen & Greenlee, 1999).

In Chapter 2, we compared VSTM for two different stimulus features, luminance contrast and orientation, under conditions of target location uncertainty. Our findings, as well as several prior results, suggest that VSTM for luminance contrast may be fundamentally different from VSTM for features such as stimulus orientation. First, luminance contrast is an intensity-coded variable (Albrecht & Hamilton, 1982), lacking the precisely-structured representational maps that encode for features such as orientation. Thus, luminance contrast is likely encoded into VSTM in a much more abstract way than other visual features (Xing et al., 2014). Second, evidence exists suggesting that memory for luminance contrast is impoverished: for example, contrast-discrimination thresholds increase substantially with inter-stimulus delay (Magnussen et al., 1996; Magnussen & Greenlee, 1999), and in the presence of distractor stimuli (Pestilli et al., 2011). Unfortunately, our understanding of the encoding and retention of luminance contrast information is based primarily on the results of coarse, 2-IFC discrimination tasks (Nachmias & Sansbury, 1974; Legge & Foley,

1980; Magnussen & Greenlee, 1999). While one prior investigation did use delayed estimation to study luminance contrast encoding, this study focused on manipulations of attention, and lacked the detailed parametric variation now common in delayed-estimation tasks (Prinzmetal et al., 1997).

These factors motivated us to examine VSTM for luminance contrast using the delayed-estimation paradigm. In Experiment 1, observers were instructed to hold in memory the perceived luminance contrast of a briefly flashed circular disc, and after a brief delay, to reconstruct the memorized contrast by adjusting the luminance of a subsequently presented match disc (Figure 4.1). We systematically measured estimate distributions for luminance contrasts spanning from low (7%) to high (76%), and found very consistent profiles of performance across observers. Control experiments investigated how the distribution shapes depended on the onset contrast of the match disc and on the polarity of the stimuli. We then developed a low-parameter, neurally plausible model of observer performance on the task, a model which incorporated a realistic form of contrast response function (i.e., Naka-Rushton), and assumed maximum-likelihood read-out. Our encoding-decoding model successfully described the shape of estimate distributions for individual observers, predicting neurally plausible gain parameter values.

4.2 Materials and Methods

4.2.1 Participants

Data from eight observers (one author) were collected in Experiment 1. Observers were recruited from the local community and student body at New York University (paid $10/hr), and amongst lab colleagues. Observers had varying degrees of experience in psychophysi-

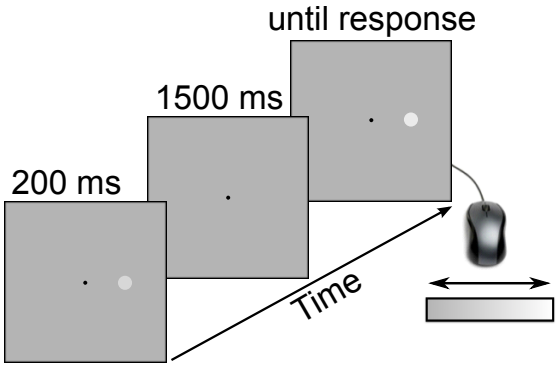

Figure 4.1: Delayed estimation of luminance contrast: experiment. Observers were briefly presented (200 ms) with a small luminance-defined, circular disc (1° diameter), either left or right (4° eccentricity) of fixation. The disc could appear with one of eight luminance contrasts on different trials. After the first disc disappeared and following a brief delay (1500 ms), a match disc appeared at the same location as the first, this time with a random luminance contrast. Observers had to estimate the luminance contrast of the first disc by adjusting (with the computer mouse) the luminance contrast of the match disc. After estimating the luminance contrast of the first disc as accurately as possible, observers recorded their estimate by pressing the left mouse button.

cal testing. All observers gave written informed consent, and experiments were carried out with approval of the NYU University Committee on Activities Involving Human Subjects.

4.2.2 Experiment 1

Task

An example trial is illustrated in Figure 4.1. On each trial of the experiment, observers were briefly presented (200 ms) with a small luminance-defined disc (1° diameter) on the computer monitor, either left or right (4° eccentricity) of the black fixation dot. On any trial, the disc appeared with one of eight luminance contrasts (see Test set-up and stimulus design), which was selected from a randomly-shuffled array within each block. After the first disc disappeared and following a brief delay (1500 ms), a second 'match' disc appeared at the same location as the first, this time with a luminance contrast chosen randomly from the range of thirty-eight possible estimate luminance contrasts (see Test set-up and stimu-

lus design). Observers were required to estimate the luminance contrast of the first disc by adjusting (with the computer mouse) the luminance contrast of the match disc. Luminance contrast of the match disc was adjusted by making small horizontal motions of the mouse - leftward motions of the mouse made the disc appear dimmer, rightward motions of the mouse made the disc appear brighter. Observers were instructed that there was no set interval for responding, and to try to perform as accurately as possible. After estimating the luminance contrast of the first disc as accurately as possible, observers pressed the left mouse button to record their estimate.

Observers completed four test sessions, of about 1 hr duration each. Each session consisted of five 80-trial blocks (10 trials per luminance contrast level), preceded by one 40-trial practice block. This gave a total of 1600 test trials per observer (i.e., 200 trials per luminance contrast). During each block, observers rested their chin on a chin-rest, and were instructed to maintain fixation on the central fixation dot throughout each trial (during presentation of the first disc and while adjusting the match disc). Observers also received limited motivational feedback, after every second block (e.g., 'Well done! You are performing above average.' or 'Good job. Your performance level is around the median of all observers.'). The feedback statements above were alternately selected at random, and feedback was not related to any performance criterion per se.

Test set-up and stimulus design

Stimuli were presented in a darkened room on an iPad retina display (monitor only), controlled by a Windows-based PC running MATLAB (The Mathworks) and the Psychophysics Toolbox. Resolution of this small monitor was 2048 pixels x 1536 pixels. The display was

controlled by an AbuseMarK LCD adapter and fixed in a custom-frame affording three degrees of freedom in monitor positioning. Before beginning each session, the experimenter ensured that the monitor was positioned centrally in front of the observer, with the fixation dot at eye height. The display was positioned in landscape mode (i.e., with the higher resolution along the horizontal). Viewing distance was set to 28.5 cm.

Monitor brightness was maintained at its default maximum setting (which allowed for maximum luminance values up to 390-400 cd/m^2). To control gray levels appropriately, we first manually gamma-corrected the display across its full range of luminance output (using a Spectrascan PR650 photometer with uniform luminance across the monitor). We then created a reduced luminance range look-up table spanning one-quarter of the full luminance range (i.e., from 0-93.5 cd/m^2), by selecting the first quarter of the gamma-corrected, full range look-up table and interpolating intermediate values to create a vector of 256 RGB intensity values. This reduced luminance range is comparable to ranges typically reported in studies where luminance contrast is manipulated (i.e., in CRT-based experiments). The reduced range, however, meant that nearby RGB indices overlapped somewhat in output luminance. Using the photometer, we manually measured the actual output luminance (with several repeats) for each RGB index in the reduced range. For the positive luminance deflections used in Experiment 1, the 129 RGB levels used (i.e., the background index of 127 and the 128 levels above) mapped onto thirty-eight unique luminance output values, which we used to naturally bin observers' estimates based on the mouse position-RGB index mapping.

We defined the luminance contrast of these thirty-eight unique levels in terms of Weber

contrast,

$$c_{\text{weber}} = \frac{I_{\text{stimulus}} - I_{\text{background}}}{I_{\text{background}}} \tag{4.1}$$

where I_{stimulus} was the disc luminance and $I_{\text{background}}$ was the luminance of the gray background. The eight test luminance contrasts were set according to fixed RGB indices using a pseudo-linear increment array; specifically, RGB values of 127 + [8, 16, 24, 32, 48, 64, 80, 96] were used. Using the measured luminance values for these indices, the resulting luminance contrasts spanned from approximately 7-76% contrast. The circular discs measured 1° in diameter, and were uniform in luminance, except for the very edge of the disc (raised-cosine, edge width 0.1°).

4.2.3 Experiments 2 and 3

We ran two control experiments (n = 8 observers each), in an effort to understand the possible roles played by a number of task and stimulus-related factors in Experiment 1. Participant recruitment and informed consent procedures followed similar protocols to those described above.

Experiment 2 was identical to Experiment 1, except that the match disc was set to 0% contrast at onset. Experiment 3 was identical to Experiment 1 except that dark luminance discs were presented on a gray background. We subtracted the array of RGB increments from mid-level gray to calculate each of the eight tested luminance contrasts. We used the same restricted gamma table (0-93.5 cd/m^2). For each RGB index below mid-level gray, we measured the actual output luminance value using the photometer, and confirmed that there were 76 unique luminance values in the 0-127 RGB range. Note that the eight luminance contrasts tested in Experiment 3 differed by small percentage amounts from those used in

Experiment 1, presumably due to limitations of the hardware and the LUT discretization described above.

4.2.4 Summary data analyses

We calculated the median and inter-quartile range of each distribution for each observer separately. We then fit the interquartile range data with a power law, using the MATLAB function *lsqcurvefit.m*, to ascertain whether the data conformed approximately to Weber's Law. That is, we found the least-squares fit that best described the data according to:

$$\hat{c}_{75} - \hat{c}_{25} = kc^w \tag{4.2}$$

where k scales the power law, with exponent w, relating luminance contrast of the disc to the width (interquartile range) of the estimate distribution. Perfect Weber's Law behavior would give an exponent of one for this relation; in investigations of luminance contrast discrimination, best-fit slope values in the range of 0.5-0.7 have been typically reported (Legge & Foley, 1980; Pestilli et al., 2011). In the main text, we report the mean (s.e.m.) of the fitted slope to two decimal places.

4.2.5 Probabilistic model of neural responses

Maximum-likelihood estimation of contrast

We begin by describing the generative model for our task (Figure 4.2). A stimulus of luminance contrast c is presented to the observer. The stimulus is encoded by a population of noisy sensory neurons, giving rise to a vector of spike counts r = $\{r_1, r_2, r_3, ..., r_n\}$. We as-

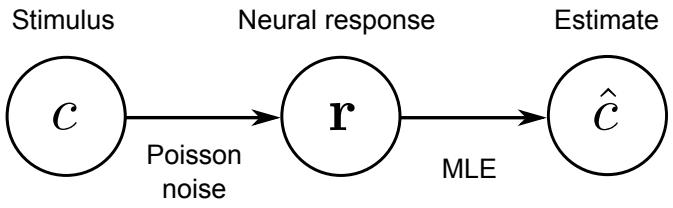

Figure 4.2: Delayed estimation of luminance contrast: encoding-decoding model. We assume that the presented contrast, c, gives rise to normally-distributed, Poisson firing rates in an array of neurons, r, and that the observer performs maximum-likelihood estimation on these responses, to obtain the estimate $\hat{c}$.

sume that the observer then decodes this set of spike counts using maximum-likelihood estimation, to arrive at an estimate $\hat{c}$.

We assume that spike counts are independent across neurons and governed by Poisson noise. The probability of spike count vector r, given luminance contrast c, is thus (Dayan & Abbott, 2001),

$$p(\mathrm{r}|c) = \prod_i \frac{1}{r_i!} e^{-g_i(c)} g_i(c)^{r_i}, \tag{4.3}$$

where $g_i(c)$ represents the contrast gain function for neuron i. We assume this gain function takes the form of a Naka-Rushton equation (Naka & Rushton, 1966),

$$g_i(c) = a_i \frac{c^n}{c^n + c_{50}^n} \tag{4.4}$$

where the responsiveness to contrast is governed by the exponent, n, the semi-saturation contrast, c_{50}, and the maximum firing rate for the neuron, a_i. We assume that maximum firing rates can vary across neurons, but that they might average out in the population during read-out. For simplicity, we did not consider heterogeneity in the other sensitivity parameters. Assuming the observer is performing maximum-likelihood estimation on the underly-

ing firing rates, then the estimate contrast, $\hat{c}$, is given by

$$\hat{c} = \arg\max_{c} \log p(\mathrm{r}|c) \tag{4.5}$$

Some straightforward calculations show that,

$$\begin{aligned} \hat{c} &= \arg\max_{c} \Big(-g(c) \sum_i a_i + \log g(c) \sum_i r_i \Big) && (4.6) \\ &= g^{-1}\left(\frac{r}{a}\right), && (4.7) \end{aligned}$$

where $a = \sum_i a_i$, $r = \sum_i r_i$, and g^{-1} is the inverse function of g. If we approximate the Poisson distribution by a normal distribution, then $r \sim \mathcal{N}(ag(c), ag(c))$ and

$$\frac{r}{a} \sim \mathcal{N}\left(g(c), \frac{g(c)}{a}\right). \tag{4.8}$$

By transforming this probability distribution under the mapping $\frac{r}{a} \mapsto g^{-1}\left(\frac{r}{a}\right)$, we obtain the conditional probability of estimate contrast, $\hat{c}$, given c, as

$$p(\hat{c}|c) = \sqrt{\frac{a}{2\pi}} \frac{g'(\hat{c})}{\sqrt{g(c)}} e^{-\frac{a}{2}\frac{(g(\hat{c})-g(c))^2}{g(c)}} \tag{4.9}$$

Model fitting

Using a maximum-likelihood procedure, we found the best-fitting parameter values of the model for each observer individually. In model fitting, we also allowed for a lapse rate parameter, λ. We assume that lapses were uniformly distributed across possible estimate values. Thus, there were four free parameters in total - three gain parameters (a, n, c_{50}) and λ.

Model-fitting was done throughout using the MATLAB function *fminsearch.m*, which is a standard optimization algorithm based on the Nelder-Mead simplex algorithm, and suitable considering the relatively small number of free parameters. We allowed each of the parameters to vary within broad but finite bounds (see Table 4.1).

4.3 Results

4.3.1 Delayed estimation of luminance contrast

Observers ($n = 8$) were presented with small, briefly flashed (200 ms) circular discs in Experiment 1, and had to reconstruct the presented contrast by adjusting the luminance of a match disc (via horizontal movements of a computer mouse). The median and inter-quartile range of estimate distributions for several individual observers are illustrated in Figure 4.3A, and example estimate histograms from Experiment 1 are depicted in Figure 4.4A. In general, the position and shape of the distributions changed in a highly consistent fashion across all observers tested. As luminance contrast of the disc increased, so too did the median and width of the estimate distributions. There appeared to be a systematic tendency for the median estimates to be shifted slightly towards the mean presented luminance contrast; this effect could have several possible causes, such as the bounded nature of the response range, some form of effort-versus-accuracy trade-off in adjusting the match disc luminance, or a Bayesian prior.

The dependence of the estimate inter-quartile range on disc contrast was well fit by a power law, with the mean (s.e.m.) of the exponent across observers equal to 0.54 (0.05) (Figure 4.3B). Thus, the exponent, or slope on a log-log axis, stood at a value similar to those typically found in traditional contrast-discrimination tasks (Legge & Foley, 1980;

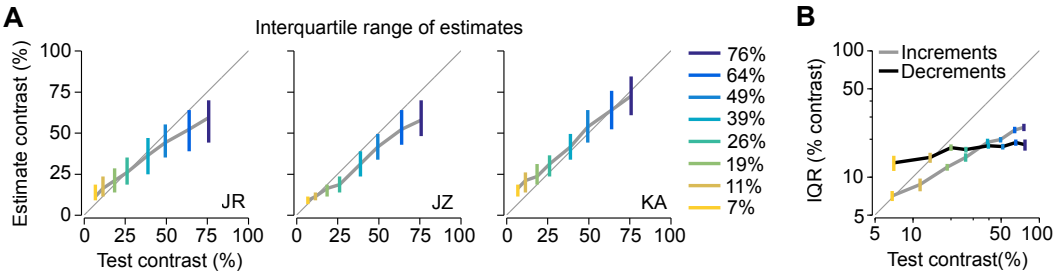

Figure 4.3: Delayed estimation of luminance contrast: data. A) Results from Experiment 1 are illustrated, with plots of the median and inter-quartile range of estimates for several individual observers. The circular disc could take one of eight luminance contrasts (different colors) across trials. B) Mean (s.e.m.) of the inter-quartile range across observers is plotted against test luminance contrast, for Experiment 1 (increments) and Experiment 3 (decrements). For positive luminance increments, the data conformed to a near-miss to Weber's Law, while the slope of the function was much more flat for decrements. Note that, for Experiment 3, several of the test contrasts differed slightly from the color-coded percentage values given in A); for clarity of presentation, we utilize the same color-coding scheme for both datasets.

Pestilli et al., 2011). For example, Legge & Foley (1980) found the slope of the relation between pedestal contrast and contrast-discrimination threshold to lie at around 0.6, a relation commonly referred to as the near-miss to Weber's Law (May & Solomon, 2015).

4.3.2 Experiments 2 and 3

In a pair of control experiments, we tested whether these results were dependent on the match disc having a non-zero onset contrast, and on the polarity of the stimuli. As onset contrast was selected randomly in Experiment 1, the match disc was more often than not of higher contrast than a recently-presented low-contrast test disc. Thus, some local adaptation or memory substitution process could have systematically affected the shape of estimate distributions for low contrast stimuli. By starting the match at 0% contrast, and requiring observers to 'dial up' the memorized target contrast, we hoped to ascertain whether such effects might be present in the data. Data collected in this first control experiment were largely indistinguishable from results of Experiment 1, with the relationship between

inter-quartile range and test contrast again following a near-miss to Weber's Law i.e., with a mean (s.e.m.) slope on a log-log axis of 0.66 (0.07). It is possible that the slightly steeper slope compared to Experiment 1 reflects some systematic process; for the present purposes, however, we see no reason to question the general method of probing used in Experiment 1.

We also tested the role of disc polarity in our paradigm. In an idealized scenario, we had assumed that disc polarity would not matter greatly. However, there is some evidence to suggest that dark and light patches play asymmetric roles in luminance and contrast discrimination at high luminance values (Whittle, 1986; Kingdom & Whittle, 1996), as well as evidence for a general asymmetry in the neural representations of darks and lights (Yeh et al., 2009; Kremkow et al., 2014). To test this possibility, we re-ran the basic experiment, this time using negative luminance increments instead of positive. Of note, many observers in this task now had a substantially broader estimate distribution for low-contrast stimuli, with the distributions increasing in width only negligibly with increasing contrast (Figure 4.3B, and see raw data Figure 4.4B). On a log-log axis, the slope relating test contrast and inter-quartile range was much shallower than for the increment experiments, with a mean (s.e.m.) across observers of 0.15 (0.05).

Is there some asymmetry in luminance processing at play here? For the present moment, we can only speculate on what the important factors are. For example, there is some evidence to suggest that positive and negative luminances give rise to more quickly saturating or more linear neural responses, respectively (Kremkow et al., 2014). Perhaps the relatively broad distribution of estimates for low-contrast decrements, relative to increments, reflects an asymmetry in the slope of the initial part of the contrast response, an effect which might arise potentially very early in visual processing (Kremkow et al., 2014). In passing,

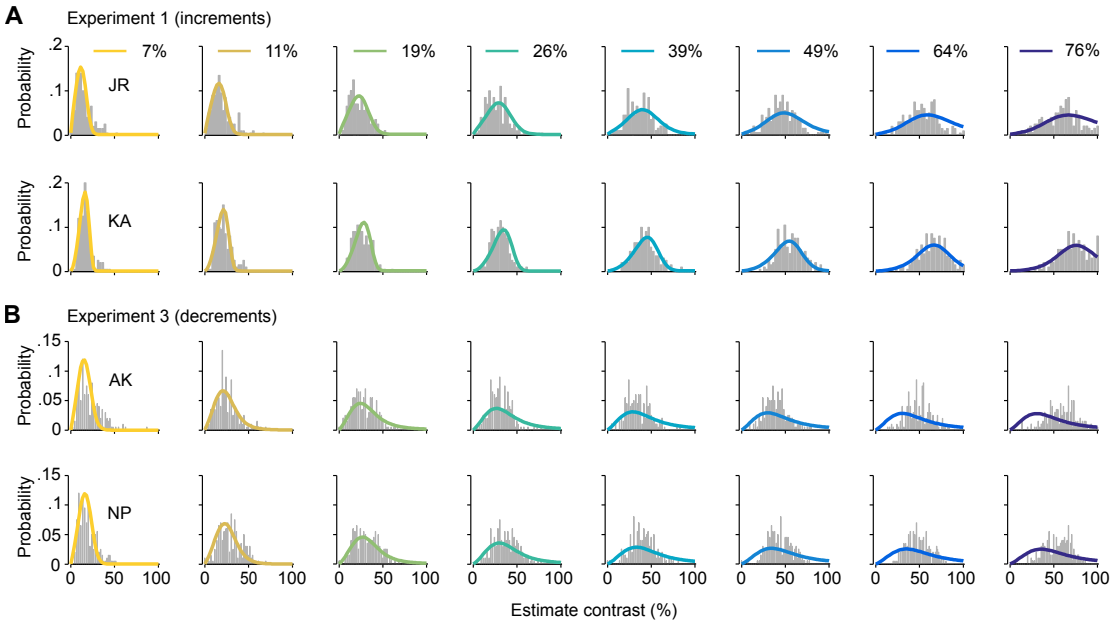

Figure 4.4: Delayed estimation of luminance contrast: model fits. A) and B) Fits of the probabilistic model to estimate distributions from Experiment 1 (increments) and Experiment 3 (decrements), respectively. Example data from a couple of individual observers are presented for each experiment. Note that, for Experiment 3, several of the test contrasts differed slightly from the color-coded percentage values given in A); for clarity of presentation, we utilize the same color-coding scheme for both datasets. In addition, note that the discretization of the estimate response axis differed across experiments, reflected here in differences in the granularity of the histogram bars across experiment. See Materials and Methods for details.

it is worth mentioning that there is some behavioral evidence to suggest that positive and negative luminance processing asymmetries underlie an effect we discussed in Chapter 3, specifically the decrease in contrast-discrimination thresholds for high-contrast sinusoidal gratings (Whittle, 1986; Kingdom & Whittle, 1996).

4.3.3 Model

We developed a four-parameter model of contrast estimation based on a hypothesized neural substrate (Figure 4.2). Specifically, we assumed that luminance contrast, c, is encoded by

a population of noisy Poisson neurons whose mean spike counts are related to c via a Naka-Rushton gain function. If we assume that the observer performs maximum-likelihood estimation on the responses to obtain a contrast estimate $\hat{c}$, then a simple, closed-from expression for the conditional probability distribution, $p(\hat{c}|c)$, can be derived (see Materials and Methods).

The model provided quantitatively good fits to the estimate distributions from individual observers in Experiment 1, with the widths of the predicted distributions systematically increasing as a function of test luminance contrast (Figure 4.4A). Parameter estimates from fits to individual observer data are provided in Table 4.1. The exponent of the Naka-Rushton equation lay consistently in the range around 2, with a mean (s.e.m.) across observers of 2.06 (0.21). Values in this range have been consistently observed in luminance- and contrast-response measurements in early visual cortical areas, with smaller values typically reported in fits to retinal data (Wilson, 1999). The preponderance of fitted values near 2 is also interesting given the suggested importance of this range of exponent value for information transmission efficiency in the contrast response (Gottschalk, 2002).

There did appear, however, to be some systematic deviations between the measured and fitted distributions. For example, model fits to data from the reversed polarity experiment (Experiment 3) were, on visual inspection, less satisfactory in general (Figure 4.4B), suggesting again that some modified contrast response model might be appropriate when fitting that dataset. In addition, for the low-contrast increment conditions in Experiment 1, the model seemed to over-estimate the position of the peak of observer estimates, with the actual estimates weighted towards lower contrasts (Figure 4.4A). This effect was present in numerous individual observers data, and is puzzling considering that the median estimates

Delayed estimation of luminance contrast: model parameters

	S1	S2	S3	S4	S5	S6	S7	S8	Mean (s.e.m.)
Param									
a	10.1	16.6	9.3	9.7	9.4	30.7	15.1	11.3	14.0 (2.6)
n	1.79	1.85	2.74	2.11	2.75	0.88	2.40	1.96	2.06 (0.21)
c_{50}	0.89	0.47	1.00	0.44	0.73	0.14	0.98	1.00	0.71 (0.11)
λ	0.065	0.002	0.055	0.039	0.044	0.0001	0.034	0.037	0.035 (0.008)

Table 4.1: Best-fit parameter values of the encoding-decoding model fits to Experiment 1 data. Parameters could vary within broad but finite bounds: a, 1-50; n, 0.1-4; c_{50}, 0-1; λ, 0-1. In the table, fitted values and their mean (s.e.m.) are rounded, with a level of precision that seemed appropriate for each parameter type.

were biased in the opposite direction. A number of possible factors might be involved. First, our assumption that all neurons have the same $g(c)$ is unrealistic (Albrecht & Hamilton, 1982). Second, we used maximum-likelihood instead of posterior mean read-out; one could imagine that in an estimation task, observers minimize the expected squared error and therefore choose the posterior mean. Third, the normal approximation to Poisson firing statistics might not be adequate; indeed our assumption of perfectly Poisson noise may be insufficient in general, with a potential role played by trial-to-trial gain fluctuations in defining the shape of the estimate distributions (Goris et al., 2014; May & Solomon, 2015).

4.4 Discussion

By providing continuous, high-resolution measurements of memory contents, delayed-estimation tasks have elucidated the nature of VSTM, most notably for orientation and color (Zhang & Luck, 2008; Bays et al., 2009; van den Berg et al., 2012). Such features are relatively stable in memory across time, presumably due to the topographic form of their

neural representations. For intensity-coded features such as luminance contrast, however, VSTM might be less stable over time (Magnussen & Greenlee, 1999). Unfortunately, the processes of encoding and retaining luminance contrast information over brief delays have been characterized predominantly using relatively coarse, 2-IFC discrimination tasks (Legge & Foley, 1980; Pestilli et al., 2011). Here, we examined memory for luminance contrast using delayed estimation. We systematically measured estimate distributions for luminance contrasts spanning the contrast axis, using small uniform discs as stimuli. Memoranda for specific luminance contrasts were clearly well-defined, with estimate distributions systematically shifting position as a function of stimulus luminance contrast, and showing a small bias towards the mean presented contrast. We also found evidence of a monotonic increase in estimate distribution width as contrast increased, reminiscent of the near-miss to Weber's Law often cited in the contrast-discrimination literature (Legge & Foley, 1980; May & Solomon, 2015).

We then fit a low-parameter, neurally plausible probabilistic model to the distributions. The model assumed Poisson noise and maximum-likelihood read-out, and incorporated a realistic form of contrast response function (i.e., Naka-Rushton). Using a mixture model approach that accommodated sources of trial-to-trial noise such as lapses, we successfully fit the general shape of the estimate distributions. The model replicated the monotonic increase in estimate distribution width with increasing stimulus contrast, and predicted neurally plausible gain parameter values. Control experiments indicated that match onset contrast played no substantial role in affecting the shape of observers' estimate distributions; however, polarity of the disc appeared to be influential, with a much flatter form to the curve depicting distribution width as a function of stimulus contrast. We hypothesize

that this difference may be related to asymmetries in the neural representation of darks and lights (Whittle, 1986; Yeh et al., 2009; Kremkow et al., 2014).

Before concluding, it is appropriate that we also address several shortcomings in the present work. From an empirical point of view, one weakness of the psychophysical methods in this chapter was the lack of any control for observer eye movements. For example, observers may have systematically deviated from fixation during trials, potentially placing the small disc at different eccentricities on different trials. This could be particularly problematic for our task, given the well known differences in contrast sensitivity as a function of retinal eccentricity (Virsu & Rovamo, 1979; Regan & Beverley, 1983). Future extensions of this work should control for small eye movements e.g., by recording observer eye position during test sessions and, at analysis stage, including trial exclusion criteria similar to those used in Chapter 2. From a theoretical perspective, our model attempts might also benefit from incorporating trial-to-trial fluctuations in gain magnitude into the model. Such modulations are known to affect firing-rate statistics in systematic ways (Goris et al., 2014). It remains to be seen whether such a model would better approximate the present data.

Overall, the delayed-estimation protocol we developed is a successful first step in understanding the underlying nature of VSTM for luminance contrast. Luminance contrast encoding has typically been investigated using the discrimination paradigm, where an observer is required to detect the occurrence of a stimulus change (e.g., a contrast increment) between two temporally-separated stimuli. Performance on such tasks is typically summarized by fitting some quantitative model to the data (i.e., a psychometric function), from which a criterion performance level is read off (i.e., a threshold). By comparing thresholds across different stimulation conditions (e.g., with or without covert attention, distractors,

etc.), investigators have often sought to better understand the underlying neural processes involved in encoding and VSTM. However, whereas 2-IFC discrimination tasks only investigate the underlying probabilistic representations somewhat superficially (Magnussen et al., 1996; Pestilli et al., 2011), our simple task provided a very precise depiction of a fundamental sensory coding ability.

5 Conclusion

5.1 Encoding-decoding models of luminance contrast processing

The encoding of local stimulus properties such as orientation, color, and luminance contrast has been studied in great detail over the years, using myriad behavioral and neural recording techniques (Hubel & Wiesel, 1962; Legge & Foley, 1980; Lennie et al., 1990; Brouwer & Heeger, 2009). The starting point for such investigations has typically been at the level of single-stimulus processing, and for performance on simplified tasks such as 2-IFC discrimination. The present thesis focused primarily on the encoding-decoding of luminance contrast, a stimulus property fundamental to all of visual processing. We studied

observer behavior using two complementary experimental protocols (discrimination and delayed estimation), and for both single-stimulus and two-stimulus tasks. We also developed neural models of observer performance on these tasks. Below, we recap briefly on the main thesis results, and discuss implications of the results for our understanding of a number of topics related to sensory processing. Specifically, we first focus on the implications for research on attentional selection and VSTM, along the way discussing possible future extensions of the experimental work described in Chapters 2 and 4. We then describe potential links to research on neural noise statistics and encoding-decoding models, suggesting ways in which the encoding-decoding approaches in Chapters 3 and 4 might be extended in future investigations. We then conclude with a few brief closing statements.

5.2 Implications for research on attentional selection

Despite the central role attentional orienting plays in behavior, the neural bases of attentional modulation and selection remain poorly understood. A number of studies now illustrate that a primary neural correlate of attention consists of an additive baseline offset in neural response (Buracas & Boynton, 2007; Murray, 2008; Pestilli et al., 2011; Chen & Seidemann, 2012). For example, by simultaneously measuring behavioral performance and the BOLD fMRI response during a contrast-discrimination task, Pestilli et al. (2011) found that the enhancement in behavioral performance due to attention could be modeled by combining an additive offset in sensory response with a max-pooling rule prior to decision. Yet, a number of other findings have also emphasized the apparent multiplicative nature of attentional modulation of neural responses (McAdams & Maunsell, 1999; Reynolds & Heeger, 2009; Herrmann et al., 2010; Itthipuripat et al., 2014). In one study, the authors systemati-

cally manipulated the spatial extent of an observer's attentional focus, while measuring the contrast-dependence of orientation discrimination for stimuli of different size (Herrmann et al., 2010). Results were in general agreement with a normalization model of attention, in which attention is implemented as a multiplicative weighting of incoming sensory signals (Reynolds & Heeger, 2009).

The results of Chapters 2 provide some insight into the neural and computational processes governing attentional selection, by highlighting a difference in the efficiency of selection as a function of task: under conditions of target location uncertainty, contrast-discrimination performance at a target location was more substantially hindered by high-contrast distractors than was orientation discrimination. These results indicate that selection of sensory responses in the contrast-discrimination task was spatially coarse in nature, echoing several prior related findings (Pestilli et al., 2011; Chen & Seidemann, 2012). In contrast, orientation-related information appeared to be more efficiently selected and decoded. Arguably, the task-dependence of these effects suggests some re-examination of key ideas on the attentional modulation and selection of early sensory neural responses. In Chapter 3, we attempted to develop computational models of the contrast- and orientation-discrimination behaviors measured in Chapter 2, and found that neither of two standard sensory interaction models could convincingly replicate data from the two tasks simultaneously. Together, these empirical and computational results suggest the particular task that an observer is engaged in (e.g., contrast- vs. orientation-based) is likely also a key factor in the types of behavioral effects researchers observe in attentional selection tasks. Thus, future attempts at discriminating between additive and multiplicative effects of attention on neural response should at least acknowledge the role that feature dimension likely plays,

designing multi-dimensional sets of experiments and associated models, while keeping other attentional and stimulus manipulations constant. Some efforts in this regard have recently been made in VSTM research and elsewhere (Matthey et al., 2015; Orhan & Ma, 2015), where we now turn.

5.3 Implications for the study of VSTM

Traditionally, studies of VSTM have quantified decoding performance using relatively coarse metrics such as discrimination thresholds or change detection performance (Palmer, 1990; Magnussen & Greenlee, 1999; Ma et al., 2014). In Chapter 4, we added to a growing body of research that attempts to measure more directly the noise properties of memoranda supporting basic visual feature discrimination and comparison over brief delays (Wilken & Ma, 2004; Zhang & Luck, 2008; Fougnie et al., 2012; van den Berg et al., 2012; Bays, 2014). Using a delayed-estimation protocol, we successfully measured and characterized the shape of observers' estimate distributions for luminance contrast.

How do these results contribute to research on VSTM, and how might they be extended in future? First, we note that the recent surge of interest in using delayed estimation has almost entirely focused on the encoding-decoding of stimulus features such as orientation and color (Fougnie et al., 2012; van den Berg et al., 2012; Bays, 2014). Luminance contrast, on the other hand, has often been treated as a nuisance parameter in such investigations, or has been utilized to create coarsely defined reliability conditions e.g., low vs. high reliability. Thus, we have developed a high-quality, yet simple experimental protocol for studying delayed estimation along an intensity-coded feature dimension. The experimental results, and associated neural model in particular, help to dispel any notion that memory for luminance

contrast is not easily characterized.

Future research might develop along a number of directions, for which our protocol could serve as a basis. For example, the task in Chapter 4 could be extended to study the effects of temporal delay on VSTM for luminance contrast, thereby obtaining a more fine-grained understanding of delay effects than provided by discrimination tasks (Magnussen & Greenlee, 1999). In addition, increasing evidence suggests that individual item representations systematically decrease in precision with increasing set-size, results which add to the view that VSTM relies on a noisy, continuous neural resource (Fougnie et al., 2012; van den Berg et al., 2012). Our protocol could be extended to parametrically vary set-size, thereby providing potentially greater insight into the stimulus interactions reported in Chapter 2 and elsewhere (Pestilli et al., 2011). One obstacle prevented this type of investigation until now: for features such as orientation and color, an implicit assumption is often made that estimate distribution shape does not vary greatly along the axis of the relevant feature dimension, leading to relatively simplified designs for set-size type experiments (e.g., presenting an array of randomly oriented gabors). However, as the data in Chapter 4 illustrate, this assumption would be grossly invalid for the case of luminance contrast; estimate distributions change shape dramatically as a function of stimulus contrast. Future contrast estimation experiments that manipulate set-size would need to account for this in their design, parametrically varying the array of test contrasts and controlling for direct effects of distractors on target stimulus encoding and decoding. The models of sensory interaction described in Chapter 3 may provide some guidance here.

5.4 Neural noise and encoding-decoding

The present results may also be of relevance to recent debate on the nature of neural noise statistics and encoding-decoding performance. Numerous recent delayed-estimation tasks have consistently found that error distributions are non-Gaussian in form, perhaps reflecting trial-to-trial fluctuations in encoding precision (Fougnie et al., 2012; van den Berg et al., 2012; Bays, 2014). While the model presented in Chapter 4 did not incorporate sources of trial-to-trial gain fluctuation, more realistic noise models may be worth investigating in fitting estimation data for luminance contrast. As an intensity-coded feature, luminance contrast is likely encoded into memory in a relatively abstract, albeit firing-rate dependent way (Albrecht & Hamilton, 1982; Xing et al., 2014). Thus, trial-to-trial fluctuations in gain magnitude (i.e., double stochasticity) would presumably directly affect the trial-to-trial variation in estimates made for a given luminance contrast. Overall, estimate distibutions for luminance contrast might be impacted more directly by the stochastic properties of early sensory encoding i.e., the shape of observers' estimate distributions might reflect properties of Poisson or super-Poisson noise statistics in some principled fashion, with predictable variation across different luminance contrast levels (Shadlen & Newsome, 1998; Goris et al., 2014). In contrast, the key limiting factors on precision for tasks involving circular, Gaussian-like tuning functions (e.g., orientation tuning curves) might be network computations that affect response amplitudes for all simultaneously-stored items (e.g., divisive normalization), as well as noise fluctuations that resemble the signal of interest (Bays, 2014; Moreno-Bote et al., 2014). It remains to be seen whether a model that incorporates trial-to-trial gain fluctuations will better approximate the data in Chapter 4.

5.5 Final comments

Using a combined empirical and computational approach, this thesis explored the nature of luminance contrast encoding and decoding, fundamental operations of visual system processing. By measuring observer performance in a variety of behavioral tasks, and fitting appropriately chosen mathematical models to the data, we highlighted important characteristics of the encoding and decoding of stimulus luminance contrast, such as the large effect of irrelevant distractors on basic discrimination abilities, and the profile of observers' internal, noisy estimates of luminance contrast. The thesis findings are relevant to a variety of subfields within the visual and sensory neurosciences, such as research on attention, memory and general models of stimulus encoding-decoding.

References

Abbott, L. F. & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation*, 11, 91–101.

Albrecht, D. G. & Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*, 48, 217–237.

Averbeck, B. B. & Lee, D. (2006). Effects of noise correlations on information encoding and decoding. *Journal of Neurophysiology*, 95, 3633–3644.

Awh, E. & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5, 119–126.

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *The Journal of Neuroscience*, 34, 3632–3645.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9(10)(7), 1–11.

Bays, P. M. & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321, 851–854.

Berens, P., Ecker, A. S., Cotton, R. J., Ma, W. J., Bethge, M., & Tolias, A. S. (2012). A fast and simple population code for orientation in primate V1. *The Journal of Neuroscience*, 32, 10618–10626.

Blake, R. & Holopigian, K. (1985). Orientation selectivity in cats and humans assessed by masking. *Vision Research*, 25, 1459–1467.

Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision Research*, 39, 257–269.

Bradley, A. & Ohzawa, I. (1986). A comparison of contrast detection and discrimination. *Vision Research*, 26, 991–997.

Brouwer, G. J. & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *The Journal of Neuroscience*, 29, 13992–14003.

Buracas, G. T. & Boynton, G. M. (2007). The effect of spatial attention on contrast response functions in human visual cortex. *The Journal of Neuroscience*, 27, 93–97.

Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of concurrent stimuli by population activity in visual cortex. *Neuron*, 64, 931–942.

Callaway, E. M. (2003). Cell types and local circuits in primary visual cortex of the macaque monkey. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 680–694). Cambridge: MIT Press.

Carandini, M. & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13, 51–62.

Chen, Y. & Seidemann, E. (2012). Attentional modulations related to spatial gating but not to allocation of limited resources in primate V1. *Neuron*, 74, 557–566.

Chirimuuta, M. & Tolhurst, D. J. (2005). Does a Bayesian model of V1 contrast coding offer a neurophysiological account of human contrast discrimination? *Vision Research*, 45, 2943–2959.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research*, 44, 3223–3232.

Cover, T. M. & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley, 1st edition.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114.

Dayan, P. & Abbott, L. F. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Cambridge: MIT Press, 1st edition.

Ecker, A. S., Berens, P., Tolias, A. S., & Bethge, M. (2011). The effect of noise correlations in populations of diversely tuned neurons. *The Journal of Neuroscience*, 31, 14272–14283.

Engbert, R. & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43, 1035–1045.

Ferster, D. (2003). Assembly of receptive fields in primary visual cortex. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 695–703). Cambridge: MIT Press.

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications*, 3, 1229.

Ganguli, D. & Simoncelli, E. P. (2014). Efficient sensory encoding and bayesian inference with heterogeneous neural populations. *Neural Computation*, 26, 2103–2134.

Gegenfurtner, K. R. & Kiper, D. C. (2003). The processing of color in extrastriate cortex. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 1017–1028). Cambridge: MIT Press.

Georgeson, M. A. & Sullivan, G. D. (1975). Contrast constancy: deblurring in human vision by spatial frequency channels. *Journal of Physiology*, 252, 627–656.

Gorea, A. & Sagi, D. (2001). Disentangling signal from noise in visual contrast discrimination. *Nature Neuroscience*, 4, 1146–1150.

Goris, R. L. T., Movshon, J. A., & Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, 17, 858–865.

Gottschalk, A. (2002). Derivation of the visual contrast response function by maximizing information rate. *Neural Computation*, 14, 527–542.

Graf, A. B. A., Kohn, A., Jazayeri, M., & Movshon, J. A. (2011). Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature Neuroscience*, 14, 239–245.

Hansen, N. & Ostermeier, A. (1996). Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, (pp. 312–317).

Hara, Y. & Gardner, J. L. (2014). Encoding of graded changes in spatial specificity of prior cues in human visual cortex. *Journal of Neurophysiology*, 112, 2834–2849.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9, 181–197.

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

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–154.

Itthipuripat, S., Ester, E. F., Deering, S., & Serences, J. T. (2014). Sensory gain outperforms efficient readout mechanisms in predicting attention-related improvements in behavior. *The Journal of Neuroscience*, 34, 13384–13398.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neuroscience*, 4, 409–416.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, 59, 193–224.

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 USA*, 83, 2755–2757.

Kingdom, F. A. A. & Whittle, P. (1996). Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. *Vision Research*, 36, 817–829.

Kremkow, J., Jin, J., Komban, S. J., Wang, Y., Lashgari, R., Li, X., Jansen, M., Zaidi, Q., & Alonso, J.-M. (2014). Neuronal nonlinearity explains greater visual spatial resolution for darks than lights. *Proceedings of the National Academy of Sciences USA*, 111, 3170–3175.

Lee, B. & Harris, J. (1996). Contrast transfer characteristics of visual short-term memory. *Vision Research*, 36, 2159–2166.

Legge, G. E. & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, 70, 1458–1471.

Lennie, P., Krauskopf, J., & Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *The Journal of Neuroscience*, 10, 649–669.

Luck, S. J. & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17, 347–356.

Ma, W. J., Navalpakkam, V., Beck, J. M., van den Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nature Neuroscience*, 14, 783–790.

Ma, W. J., Shen, S., Dziugaite, G., & van den Berg, R. (2015). Requiem for the max rule? *Vision Research*, 116, 179–193.

Magnussen, S. & Greenlee, M. W. (1999). The psychophysics of perceptual memory. *Psychological Research*, 62, 81–92.

Magnussen, S., Greenlee, M. W., & Thomas, J. P. (1996). Parallel processing in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 202–212.

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

Matthey, L., Bays, P. M., & Dayan, P. (2015). A probabilistic palimpsest model of visual short-term memory. *PLoS Computational Biology*, 11(1), e1004003.

May, K. A. & Solomon, J. A. (2015). Connecting psychophysical performance to neuronal response properties I: discrimination of suprathreshold stimuli. *Journal of Vision*, 15(6)(8), 1–26.

Mazyar, H., van den Berg, R., & Ma, W. J. (2012). Does precision decrease with set size? *Journal of Vision*, 12(6)(10), 1–16.

McAdams, C. J. & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *The Journal of Neuroscience*, 19, 431–441.

McIlwain, J. T. (1996). *An introduction to the biology of vision*. Cambridge: Cambridge University Press.

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., & Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, 17, 1410–1417.

Murray, S. O. (2008). The effects of spatial attention in early human visual cortex are stimulus independent. *Journal of Vision*, 8(10)(2), 1–11.

Nachmias, J. & Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision Research*, 14, 1039–1042.

Naka, K. I. & Rushton, W. A. H. (1966). S-potentials from luminosity units in the retina of fish (cyprinidae). *Journal of Physiology*, 185, 587–599.

Nilsson, T. H. & Nelson, T. M. (1981). Delayed monochromatic hue matches indicate characteristics of visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 141–150.

Ohzawa, I., Sclar, G., & Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *Journal of Neurophysiology*, 54, 651–667.

Orhan, A. E. & Ma, W. J. (2015). Neural population coding of multiple stimuli. *The Journal of Neuroscience*, 35, 3825–3841.

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

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

Paradiso, M. A. (1988). A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biological Cybernetics*, 58, 35–49.

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, 44, 369–378.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, 2, 1508–1532.

Pestilli, F., Carrasco, M., Heeger, D. J., & Gardner, J. L. (2011). Attentional enhancement via selection and pooling of early sensory responses in human visual cortex. *Neuron*, 72, 832–846.

Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. *Annual Review of Neuroscience*, 26, 381–410.

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomonology of attention: 1. color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 261–282.

Prinzmetal, W., Nwachuku, I., Bodanski, L., Blumenfeld, L., & Shimizu, N. (1997). The phenomonology of attention. 2. brightness and contrast. *Consciousness and Cognition*, 6, 372–412.

Ratliff, C. P., Borghuis, B. G., Kao, Y.-H., Sterling, P., & Balasubramanian, V. (2010). Retina is structured to process an excess of darkness in natural scenes. *Proceedings of the National Academy of Sciences USA*, 107, 17368–17373.

Recanzone, G. H., Wurtz, R. H., & Schwarz, U. (1997). Responses of MT and MST neurons to one and two moving objects in the receptive field. *Journal of Neurophysiology*, 78, 2904–2915.

Regan, D. & Beverley, K. I. (11983). Visual fields described by contrast sensitivity, by acuity, and by relative sensitivity to different orientations. *Investigative Ophthalmology and Visual Science*, 24, 754–759.

Reynolds, J. H. & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, 61, 168–185.

Sanborn, A. N. & Dayan, P. (2011). Optimal decisions for contrast discrimination. *Journal of Vision*, 11(14)(9), 1–13.

Schiller, P. H., Sandell, J. H., & Maunsell, J. H. R. (1986). Functions of the ON and OFF channels of the visual system. *Nature*, 322, 824–825.

Sergent, C., Ruff, C. C., Barbot, A., Driver, J., & Rees, G. (2011). Top-down modulation of human early visual cortex after stimulus offset supports successful postcued report. *Journal of Cognitive Neuroscience*, 23, 1921–1934.

Seriès, P., Stocker, A. A., & Simoncelli, E. P. (2009). Is the homunculus "aware" of sensory adaptation? *Neural Computation*, 21, 3271–3304.

Seung, H. S. & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences USA*, 90, 10749–10753.

Shadlen, M. N. & Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of Neuroscience*, 18, 3870–3896.

Shapley, R. (1990). Visual sensitivity and parallel retinocortical channels. *Annual Reviews of Psychology*, 41, 635–658.

Shapley, R. & Victor, J. D. (1979). The contrast gain control of the cat retina. *Vision Research*, 19, 431–434.

Sincich, L. C. & Horton, J. C. (2005). The circuitry of V1 and V2: integration of color, form, and motion. *Annual Reviews of Neuroscience*, 28, 303–326.

Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I., & Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior. *Journal of Neurophysiology*, 57, 773–786.

Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Experimental Brain Research*, 41, 414–419.

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

Virsu, V. & Rovamo, J. (1979). Visual resolution, constrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, 37, 475–494.

Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orientation sensitivity. *Journal of the Optical Society of America*, 66, 332–338.

Whittle, P. (1986). Increments and decrements: luminance discrimination. *Vision Research*, 26, 1677–1691.

Wichmann, F. A. & Hill, N. J. (2001). The psychometric function: I. fitting, sampling, and goodness of fit. *Perception and Psychophysics*, 63, 1293–1313.

Wiesel, T. N. & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, 29, 1115–1156.

Wilken, P. & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4, 1120–1135.

Wilson, H. R. (1999). *Spikes, decisions and actions: dynamical foundations of neuroscience*. Oxford: Oxford University Press.

Xing, Y., Ledgeway, T., McGraw, P., & Schluppeck, D. (2014). The influence of spatial pattern on visual short-term memory for contrast. *Attention, Perception and Psychophysics*, 76, 1925–1932.

Yeh, C.-I., Xing, D., & Shapley, R. M. (2009). "Black" responses dominate macaque primary visual cortex V1. *The Journal of Neuroscience*, 29, 11753–11760.

Zenger-Landolt, B. & Heeger, D. J. (2003). Response suppression in V1 agrees with psychophysics of surround masking. *The Journal of Neuroscience*, 23, 6884–6893.

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

Zoccolan, D., Kouh, M., Poggio, T., & DiCarlo, J. J. (2007). Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *The Journal of Neuroscience*, 27, 12292–12307.