Temporal and spatial properties of the BOLD fMRI response to first and second order contrast in V1

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Serena Kainoa Au Thompson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Daniel Kersten and Cheryl Olman

January 2010

Acknowledgements

I would like to thank Jenny Schumacher who collaborated with me throughout every study herein, Essa Yacoub, who was instrumental in both the acquisition and my understanding of the 7 Tesla fMRI experiments, as well as my advisers and mentors: Cheryl Olman, Daniel Kersten, and Stephen Engel. Thanks especially to Cheryl, who was always available and willing to help me with whatever aspect of science or life was alluding my understanding, and my family – parents Raydeen and Galen, brother Keenan, and furry members Pixie, Charlie, and Rosie.

Abstract

Our ability to detect, discriminate, and identify objects, to extract depth information through the use of stereopsis, and to learn about and classify different surface properties and textures are all secondary to first extracting useful cues about the local contrast within a visual scene. Contrast can be made by a change in a variety of visual features. We will specifically consider the contrast made by first-order and second-order luminance changes, specifically the contrast between local pixel luminance, and the contrast between local pixel luminance variance. There are a variety of tools available to 'observe' cortical modulation to contrast, such as electrophysiological recordings of spike rate or local field potentials, fMRI BOLD modulation, magnetic encephalography (MEG), or visually evoked potentials (VEP), or through behavioral markers such as detection thresholds and discrimination thresholds. BOLD fMRI modulation provides a unique tool to measure the early visual response to local changes in image contrast with superior spatial specificity and minimal subject invasiveness, while psychophysics allows us to quantify the information human observers use to detect contrast. We will use both BOLD fMRI and psychophysics to explore three components of the human response to contrast: 1) how much do we use contrast when it forms the edge of a target shape compared to when it composes the interior area of the shape; 2) how well does second-order contrast modulate early visual cortex compared to the modulation elicited by first-order contrast; and 3) how accurate is the timing information of functional magnetic resonance imaging (fMRI) of the blood oxygen-level dependent (BOLD) modulation.

Human observers do not always use contrast in the same way, and in the first experiment we study how human observers use first-order contrast in an image region versus second-order contrast at a region boundary when accomplishing a difficult detection task. We may expect V1 response to differ when contextual cues suggest that contrast is contained within an object border compared to when it fills in the interior texture of an object. Consider the case of two Holstein (black and white spotted) cows standing together, one partially occluding the other. We can use contrast cues to both identify which parts belong to which cow, as well as to determine the type of cow based on the spotted character of its hide. Both the segmentation task and the identification task require use of the same type of contrast information. However, optimal processing may change the way that low-level cues are handled in early visual cortex by modulation through context-dependent feedback as well learned information about cows. This may give rise to similar low-level cues producing different neural signals in early cortical areas, and subsequently, different sensitivity to cues in a border compared to a region. Our first experiment explores observers' use of contrast contained within the edge of a detectable target compared to the contrast that makes up the entire interior region of the target.

Extraction of local image contrasts occurs at an early stage of visual processing, however, which types of contrast and how strongly they modulate early visual cortex remains undetermined. There are a variety of contrasts that could be compared. The first-order versus second-order contrast comparison is appealing because, with images generated by locally changing either average pixel luminance or pixel variance in white

noise, we can create stimuli that are equal in orientation and spatial frequency information, contain detectable boundaries, and still require independent information to be extracted for detection. Our second experiment quantifies the response in primary visual cortex (V1) to first- and second-order using fMRI to acquire a spatially specific edge response in human subjects.

Finally, it would be ideal to acquire both spatially and temporally specific measurements of cortical modulation in response to contrast changes. Several research groups around the world have begun using the relative timing between BOLD fMRI events to assign causal significance to the temporal order of modulation across different cortical areas using dynamic causal modeling. If, however, the timing of the modulation depends on the strength of the modulation, we could draw false conclusions from relative timing comparisons. Therefore, in our third experiment we measured three temporal characteristics of the BOLD hemodynamic response function (HRF) as a function of stimulus contrast: onset latency, time to peak, and full-width half-maximum.

These explorations into the human and BOLD fMRI response to contrast are aimed at developing the current knowledge about how we perceive and parse the world around us, as well as how we can better interpret one of our measurement tools, fMRI, as a correlate of the neuronal activity occurring in early visual cortex. Both aims guide us toward a better understanding of the mechanisms we employ to process the rich information our visual systems acquire.

Table of Contents

List of Tables

List of Figures

Chapter 1: Behavioral contrast response as a function of shape: area versus perimeter contributions

We know that segmenting objects from their backgrounds is crucial for organizing the visual scene, and that the information for object segmentation comes from two sources: edges and regions. The extent to which we use the information provided by an object's edge compared to the information contained in regions within an object that can be grouped by similarity is still debated. To further pursue this investigation, we compared contrast detection thresholds for two different shapes of stimuli (annulus- and propeller-shaped), equating for either perimeter or area. Although there are an unlimited number of possible textures in which we could embed our target shapes, we used white noise with different pixel variance inside versus outside the target shape for several reasons. White noise is rich in non-redundant pixel luminance information, and there are a variety of pixel luminance distribution characteristics we can easily change to define unique textures that are both differentiable by human observers and easily mathematically quantifiable.

We chose to define the target shapes using second-order contrast. Contrast-defined or second-order edges describe a change in luminance contrast, or variance. Whereas we define first-order contrast by Michelson contrast:

$$\frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

where I_{max} is the maximum pixel intensity and I_{min} is the minimum pixel intensity within a given area.

And second-order contrast is defined by the difference in contrasts, expressed as:

$$\frac{c_{max} - c_{min}}{c_{max} + c_{min}}$$

where c_{max} is the Michelson contrast calculated within a target area and c_{min} is the Michelson contrast calculated outside a target area.

We already know that several edge statistics (such as luminance and orientation) are very easily detectable and could plausibly bias an observer towards selective perimeter use resultant from learning through every-day visual detection tasks. Second-order contrasts, which are both less well studied than first-order contrasts, and less frequent in natural stimuli, may have less pre-conditioned search strategies than first-order contrasts.

To compare the contributions of area and perimeter, we measured psychometric functions for second-order contrast-defined annulus- and propeller-shaped stimuli. All stimuli were symmetric about the fovea. The annuli had inner radii of 0.875, 0.75, 0.625, and 0.5° with corresponding outer radii of 1.25, 1.5, 1.75, 2°. The propellers were made to match the largest annulus in its inner and outer radius, and invisible wedge-shaped occluders positioned over the vertical and horizontal meridia were adjusted in size so that either the total outer perimeter or the total target area matched each of the different sized annuli total area or perimeter. For most subjects, quantifying threshold as a function of area resulted in a u-shaped function for small-perimeter targets (13.4, 14.1, and 14.9° visual angle along the shape perimeter), for both annuli and propellers, while for an ideal observer that used all pixels equally, the threshold decreased monotonically with respect to both perimeter and area. All subjects also showed an almost flat relationship between

area and threshold with large propeller-shaped stimuli (perimeters of 15.3, 15.7, 19.1, and 23° visual angle).

1. Introduction

Segmentation

Our ability to detect, discriminate, classify, and identify objects and scenes in our visual world rests on our ability to segment the visual world into discrete regions. In the machine vision world, there are some computer algorithms that take advantage of edges (Rosenfield 1982; Haddon 1990) and there are others that take advantage of region similarity (Brown 1982). The algorithms that perform best, however, use both edge detection and region similarity grouping (Hurlbert and Poggio 1988; Haddon 1990; Zhu 1995).

There are numerous examples of how humans excel at segmentation. As evidence that region-grouping is enacted, segmentation is often robust even with significant degradation to objects' borders. For example, people can easily (although sometimes only after learning) identify objects with ambiguous or missing boundary information (as depicted by the Dalmatian dog or Dallenbach's cow - figure 1.1), see camouflaged animals and hidden pictures, and identify the folds in a piece of fabric by the change in texture. Conversely, people can perform visual tasks on images of objects represented only by their borders. These include silhouettes created by low light or back lit conditions, line drawings, and camouflaged animals or objects, which often have nearly

identical texture features on their interior and exterior. Thus, both edges and interior regions offer important information for an observer to use.

What exactly is an edge: 1st- and 2nd order edges

Edges can be defined by color, luminance, orientation, spatial frequency, or a variety of other changes to the visual scene and people use a variety of these edge definitions to segment objects (Beck 1987). However, edges are always a result of a change in some quantifiable statistic or quality of the 'stuff' in space, or a texture edge.

Many aspects of a texture affect how well it is perceived as a uniform field, or how well it can be differentiated from another texture. The texture feature, or 'texton', argument suggests that segmentation depends strongly on brightness, color and size of texture elements (Beck 1983; Beck 1987; Bergen and Adelson 1988), as well as blobs (Voorhees and Poggio 1988), oriented lines (Landy and Bergen 1991), line intersections, and line terminators to which the visual system pre-attentively responds (Caelli et al. 1978; Julesz 1981; Julesz and Papathomas 1984; Julesz 1986). The texton model of texture perception suggests that sometimes texture features (Beck 1983) motivate segmentation, and sometimes statistical differences motivate texture segmentation (Caelli et al. 1978). Although debate remains over whether first and second-order information in natural images are independent (Schofield 2000; Lindgren et al. 2008), Chubb and colleagues have shown that people can discriminate textures that share common mean and variance, but differ in the histogram distribution of mean pixel values (Chubb et al. 2004; Tyler et al. 2004). Because people can detect these very small differences in

texture, it seems plausible that the visual system extracts contrasts of first- and second-order information to optimize information use.

There are a variety of studies showing that second-order edges direct attentional and processing resources. For example, there is better change perception at texture edges (Moller and Hurlbert 1996), faster motion detection with second-order motion edges (than full field coherent motion), (Moller and Hurlbert 1996), and (it has been postulated) more accurate letter detection with border enhancement (Regan 1995).

It is well known that early visual cortex is sensitive to first-order edges, such as luminance edges and orientation-defined edges. We know through fMRI adaptation experiments that there are different cell populations in V1 that adapt to orientation, luminance, and contrast modulated gratings, suggesting that V1 does in fact separately process second-order contrast types (Larsson et al. 2006). Because edges are so pervasively used in visual tasks, it seems plausible that not only do early cortical areas extract a variety of edge contrasts, but they are used and processed differently depending on what information content (figure edge versus texture element) they carry.

Region boundaries affect region area response

Region information not only helps create continuity of surface, but also aids in contour definition. Similarity is the basis for some of the Gestalt properties which help observers group region information. After collecting a database of images segmented by people, the Gestalt properties of proximity, similarity, and region convexity were quantified to be informative but not predictive of human segmentation decisions, suggesting that these region-based computations are used (although not exclusively) to

inform segmentation (Martin 2001). There are also examples in which figure-ground segmentation that cannot be attributed to local orientation contrast in texture scenes, supporting the differential processing of boundaries and region information (Harrison and Keeble 2008).

We further know that region information is represented neurally and that this representation changes depending on context. There is evidence from 'filling-in', simultaneous lightness contrast, and figure/ground effects that we treat edges and the regions close to them uniquely.

Lightness studies suggest that people 'fill-in' luminance information within a bounded figure depending on the perceived strength of the figure edge. The neural response timing measured as a function of distance from the figure edge is faster at the retinotopic location of the edge than at the figure interior, suggesting that edge information is processed first, and neural processing of regions 'fills-in' from the edge, so that the neural timing is a function of the edge (Paradiso et al. 2006; Pereverzeva and Murray 2008)).

A similar effect to filling-in is simultaneous lightness contrast, or a brightness change induced to a figure by changing the luminance of the figure surround. Although there is disagreement (Cornelissen et al.), one study shows that our perception of the figure interior and V1 BOLD amplitude (Pereverzeva and Murray 2008) are consistent with the perception and response of a homogeneous disk. Illusions like the Craig-O'Brien-Cornsweet illusion (figure 1.2) are also evidence that what happens at an edge changes the percept of the adjacent regions. Finally, discriminability can increase when the size of

a surround increases even though the area corresponding to the receptive field of the measured cell does not change (Beck 1966).

Figure/ground studies also show that a neuron fires differently when the luminance edge within the neuron's receptive field signifies the left or right side of a bounded square (Zhou et al. 2000). Both behavioral results (Rubenstein and Sagi 1996) as well as single cell recordings in early visual cortex (V1, V2, V4) (Zipser et al. 1996; Zhou et al. 2000) indicate that response to homogeneous regions varies depending on whether the region is inside or outside a target or foreground area.

We probably use edges and region information

Mumford postulated that human vision uses a hybrid algorithm sensitive to both region homogeneity and edges. He found that observers experienced a decrease in polygon recognition ability either when noise was added to region information or when the edges were blurred (Mumford et al. 1987). Furthermore, people have also been shown to perform better than computer algorithms that contain both luminance and texture edge detectors (specifically when performance requires local texture comparisons) (Fogel I 1989).

Modulations of luminance and modulations of contrast could either be processed by the same mechanisms, or they could be processed independently(using multiple channels sensitive to either luminance changes or variance changes). Multichannel models have been proposed for the processing of luminance and contrast in a variety of other visual realms including depth and motion. Stereoacuity is preserved when there is no depth information provided by luminance cues in a sine wave grating patch, but depth

information can still be extracted from second-order changes in the sine wave grating carrier frequency. This effect cannot be explained by a model in which first- and second-order contrast area processed in the same channel (Wilcox and Hess 1997). Luminance- and contrast-defined disparity are also differentially sensitive to spatial frequency, resulting in humans achieving better detection of small luminance-defined stereoscopic disparities, and more dependence on image features size for contrast-defined depth cues. This differential sensitivity suggests separate spatial frequency specific channels (Wilcox and Hess 1995; Wilcox and Hess 1997; Wilcox and Hess 1998). Further, while stereoacuity improves with increasing first-order contrast (Halpern and Blake 1988), stereoacuity improves much less with second-order contrast (Wilcox and Hess 1998).

Evidence from spatial summation studies also suggests a two-channel model. Schofield and Georgeson suggest a two channel model in which one channel contains a full-wave rectifier and responds only to contrast modulated stimuli, whereas the other has no early non-linearity and responds only to luminance modulated stimuli. This model out-performs a single channel model with a half-wave recitifying non-linearity occurring immediately after the modulation transfer function (Schofield and Georgeson 1999). Moller and Hurlbert showed that the response to stripes of texture is not equal to the probability summation of narrower stripes extending over the same eccentricity (Moller and Hurlbert 1996), suggesting that there is more to the neural response to region than just the region.

Assuming there is luminance and contrast sensitivity, there are also several 2-channel theories of edge and region processing. One theory for segmentation (suggested by

experiments in coherent motion-induced segmentation) is that a (slow) local calculation for edge extraction is computed based on attribute contrasts, and a (fast) global calculation for region grouping is computed based on similarity measures (Moller and Hurlbert 1996). A competing hypothesis, supported by monkey EEG and BOLD evidence, suggests that texture edge detection occurs faster and therefore occurs through feed-forward mechanisms, while surface segregation occurs later and therefore due to feed-back, reverse hierarchical mechanisms (Scholte et al. 2008 and Lamme 2008). We would like to test the relative contributions of edge and region information to target detection. While this investigation will not differentiate between these two models of processing speeds, both models suggest differential sensitivity to edge and region information, and therefore, suggest that target detectability should be better predicted by one or the other.

Contrast-defined targets

In this study, we used contrast-defined boundaries in white noise stimuli to probe human observers' reliance on edge and region information in a contrast detection task. Because luminance is already well studied, and has known implications on the response of early visual cortex, we sought a more complex visual structure than luminance-defined stimuli. Conversely, natural textures are extremely complex, difficult to mathematically define, and therefore responses to natural textures are difficult to quantify. We chose white noise because it provides a simple and easy to define texture that provides a rich percept of surface when juxtaposed against a background that differs by average luminance or average variance. We sought to discover whether second-order contrast

defined edges elicit the same threshold detectability as luminance-contrast edges. The relationship between pedestal luminance contrast and contrast increment thresholds is known to follow a dipper function (Legge and Foley 1980) .However, fewer studies have been undertaken to describe contrast defined edges.

Compared to luminance defined blobs in static white noise, contrast-defined blobs have larger spatial summation areas at both the fovea and up to 10 degrees in the periphery again suggesting separate mechanisms for contrast and luminance (Sukumar 2007). Furthermore, sensitivity curves to luminance- and contrast-modulated Gaussian enveloped sine wave gratings with added static white noise have similar shapes as a function of spatial frequency and Gaussian envelope size. A dipper function describes the relationship for detection thresholds of luminance- and contrast-defined sine waves against increasing pedestal luminance and contrast (respectively) but increasing luminance pedestals do not affect contrast detection thresholds, nor do contrast pedestals differentially affect luminance detection thresholds (Schofield and Georgeson 1999). The null hypothesis, that edge-dependent processing mechanisms are shared for first-order and second-order contrast-defined target detection, would suggest that there is a critical size larger than which any target elicits the same detectability. Because contrast defined blobs appear to induce higher spatial summation than luminance blobs, we also expect second-order annuli to exhibit more dependence on area subtent than on total target perimeter.

2. Methods

Participants

Three experienced psychophysical observers and four novice observers (4 female, 3 male, ages 21-35) participated in this study. Five participants were naïve to the experimental aims, while the other two were authors. All participants had normal or corrected-to-normal visual acuity. Participation was voluntary and in compliance with the University of Minnesota Institutional Review Board guidelines.

Visual apparatus and stimuli

Visual stimuli were created using Matlab and the Psychtoolbox. Stimuli were presented on a NEC 2190UXi monitor with a resolution of 1024×768 pixels and a refresh rate of 60 Hz. The luminance of the monitor was calibrated using a photo research PR-655 photometer. The monitor had a range from 0.38 to 225.9 cd/m^2 with a mean gray luminance of 113.1 cd/m^2.

The stimuli used for this experiment contained one target, either an annulus or an annulus with invisible wedge-shaped occluders at the vertical and horizontal meridia producing four propeller-like shapes. “Occluded” image regions contained white noise with mean and variance matched to the background. We used four different sizes of annuli centered on 2° eccentricity. Annulus sizes were chosen to expand 0.25° at the outer edge, and only 0.125° at the inner edge to accommodate for magnification of the foveal visual scene on the cortical surface. The annuli had inner radii of 0.875, 0.75, 0.625, and 0.5° each with a corresponding outer radius of 1.25, 1.5, 1.75, 2°. This resulted in annuli with areas of 2.5, 5.3, 8.4, and 11.8 square degrees visual angle and perimeters of 13.4, 14.1, 14.9, and 15.7 degrees visual angle (shown in figure 1.3). Each

square degree visual angle area contained approximately 70 pixels, and each degree of visual angle on the perimeter contained about 4 pixels. The propeller-shaped stimuli were created in two sets: one to match the original annuli in area, and one to match the original annuli in perimeter. Both the area- and perimeter- matched annulus wedges were created by occluding the thickest annulus from the original whole set with full-screen wedges centered on the vertical and horizontal meridia. The area was matched to produce three propellers (matching the 3 smaller annuli). A second set of propeller stimuli were created by adjusting the width of the wedge occluders so the perimeters of the propeller stimuli were matched to the perimeters of the original annulus stimuli. The entire propeller set had areas of 1.0, 1.6, 2.2, 2.5, 2.8, 5.3, 8.3 square degrees visual angle, and perimeters of 13.4, 14.1, 14.9, 15.3, 15.7, 19.1, and 23.1 degrees visual angle (with about 70 pixels per square degree visual angle area and 4 pixels per degree visual angle along the perimeter).

Both the target shape and the pixels outside the target were drawn from a random uniform distribution of uniformly distributed white noise between 0.3 and 0.7 (on a luminance scale from 0 denoting black (0.38 cd/m^2) to 1 denoting white (225.9 cd/m^2), with 0.5 corresponding to the mean gray of 113.1 cd/m^2. Pixels in the target image regions (annuli or propellers) were drawn from a broader distribution, which was controlled by using a distribution of that varied through the experiment. The experiment started at the maximum of 90% contrast and decreased as a function of subject response according to two interleaved adaptive staircases (described in the next section). Contrast could be any one of twenty equally logarithmically spaced steps between 90% and 50% contrast, with smaller spacing at lower contrast. The background pixels were sampled

from a constant uniform distribution with a variance of 50%. Both background and foreground pixels were resampled from their respective distributions on each trial so that the entire image changed on each trial.

Data analysis and fitting

We used interleaved adaptive staircases in order to sample complete psychometric functions (Levitt 1971). One was 2-down 1-up, and one was 4-down 1-up, such that the first converged at a 70% correct performance level, and the second converged at an 84% correct performance level (Geisler and Chou 1995). Each staircase was run for 50 trials. Each subject completed three threshold staircase set repetitions for each of the stimuli (total of 10 target types). Subjects 3 and 6 completed an additional three sets of repetitions. Target type order was pseudo-randomized. Threshold data were grouped across size and type (annuli or propellers) and fit to sigmoidal psychometric functions using the psignifit toolbox version 2.5.6 for Matlab, a software package which implements the maximum-likelihood method described by Wichmann and Hill (see http://bootstrap-software.org/psignifit/, (Wichmann and Hill 2001)). Fits were estimated over the data from both staircases, from one day, and mean threshold over all (3-6 days) were reported. Fits were also estimated over the entire data set from each subject yielding one threshold estimate per subject. Because both threshold estimates looked very similar, and the threshold estimate varied very little for most subjects for most days, we are reporting only the fits described in the first method here. Example fits for subject 1 are shown in figure 1.4.

Ideal observer

An ideal observer simulation was performed to compare against human performance. The ideal observer has perfect information about which pixels belong to the target and which do not, as well as perfect information about the luminance of each pixel and the distribution statistics from which the target and background pixel values are drawn (as constrained by their spatial location).

The ideal observer needs very few samples in order to achieve perfect performance. This makes it hard to compare to human performance because the ideal observer is always operating at a ceiling. We therefore ran a test simulation to determine the appropriate down-sampling rate to mimic the percent correct in the human data. For each target, we had the ideal observer randomly sample from the target pixels, decreasing the sampling rate with each iteration of the simulation until performance reached about 70% correct, similar to the human data. I used the number of sample pixels needed for the smallest stimulus size for our performance criterion in the ideal observer simulation (~0.05% of the total screen size, yielding between 15 and 174 pixels within the target shape, depending on the target size). I then ran the same psychophysics experiment used to collect the human data allowing the ideal observer to make responses.

3. Results

Ideal observer

We would expect the ideal observer to perform at lower and lower thresholds as it is given more samples of the target, or as a function of area. As expected, the ideal observer shows a monotonic decrease in threshold with area regardless of whether the stimulus shape was a set of annuli or a set of propellers. Because this is a simple ideal observer,

there is no advantage to additional perimeter information in the absence of additional area information.

Threshold vs. Area or Perimeter

We observed an almost flat relationship between threshold and area or perimeter. However, thresholds for the annuli and perimeter-matched propellers were remarkably similar for each individual subject as well as for the average across subjects. Although the annuli were not significantly different by a matched paired t-test from either the perimeter-matched or the area-matched propellers (as would be expected in a completely flat relationship between threshold and area), the area-matched propeller thresholds were more different from the annuli thresholds than were the perimeter-matched propeller thresholds, and almost reached a significant level with a $p = 0.0502$.

When the target is small, there is a higher perimeter:area ratio. We see that threshold decreases for the first 3 sizes tested in this experiment. However, as the area continues to grow, both perimeter and area also increase, but the ratio between them diminishes. This change in relative perimeter and area information, or the greater spatial distribution of the available target may trigger a strategy change causing thresholds to decrease again despite a greater amount of available data.

Although we did not quantify this observation, the thresholds for both annuli and small (perimeter-matched) propellers appeared u-shaped for most subjects and for the average through subjects. This was not predicted by the ideal observer.

4. Discussion

Area psychometric function

We found a shallow u-shaped function of threshold on stimulus size regardless of stimulus shape. We also found an increasing failure of the human observed compared to the ideal observer at bigger areas. The observed performance indicates that there is an advantage to small stimulus detection that exceeds our expectation based solely on summation over the target area. This may be due to a restricted attentional demand, but cannot be explained by a change in eccentricity or cortical magnification (Cowey and Rolls 1974), as the pattern was similar for annuli, which were changing slightly in eccentricity, and for propellers, which were stationary with respect to eccentricity.

Our data are consistent with a mechanism in which perimeter information is used to improve thresholds for small stimuli, but is not used, or used less for the detection of large stimuli. A second-order edge detection mechanism or a Gestalt mechanism of edge continuity could focus attention by providing a spatial cue about relevant information (decreasing spatial uncertainty (Pelli 1985)) that is not available when the perimeter information is also spatially distributed. Alternatively, subjects are only using perimeter information when they are affected by spatial uncertainty. If the subjects know that they are under-sampling the available data (they perform at about the same threshold to the ideal observer, which is using 0.05% of the available pixels), they may be undersampling selectively at the same locations within the target, as the target does not move trial by trial. At large enough areas, subjects may be able to estimate their own uncertainty and restrict themselves spatially, but when targets become small, they need to update their estimate of the target location on each trial in order to maintain high performance. However, the medium-sized propeller thresholds are, on average, actually worse than the

smaller propeller thresholds. This suggests that disregarding the perimeter information is not optimal, and unlikely done consciously.

Although we found a small effect of size on threshold for small stimuli, we found the curve to flatten for large stimuli. This is not, however, unexpected and may either be because we presented our stimuli in static white noise, or because thresholds tend to flatten at large areas. 'Off-frequency looking' suggests that when a signal in embedded in noise, spatial frequency selective channels may become overwhelmed by power in the noise, and spatial frequency selectivity flattens (Burgess et al. 1997; Solomon 2000). This can also lead to a flattened relationship between stimulus duration and detectability, an advantage that is preserved in the absence of noise, or with dynamic noise (McAnany and Alexander 2009). Furthermore, we know that contrast sensitivity increases with area for targets with small areas (critical area depends on spatial frequency), but saturates and remains constant at large areas (Rovamo et al. 1993; Luntinen et al. 1995) when the target is embedded in noise.

5. Conclusion

Because for most subjects threshold follows a u-shaped pattern as a function of size, it is possible that subjects are using both perimeter information and area information for target discrimination. Because we found a similar threshold trend in both annuli and propeller-shaped stimuli, this threshold trend is more likely due to the perimeter:area ratio and less likely an artifact of the shape tested.

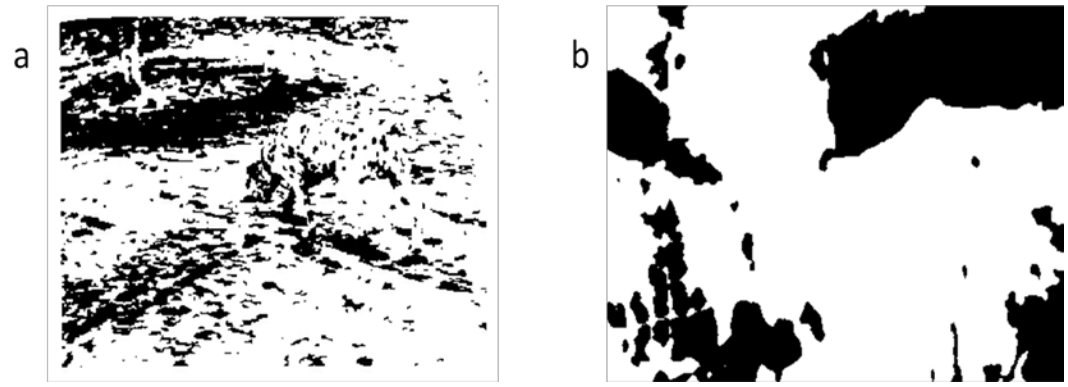

Figure 1.1: a) Dalmation b) Dallenbach's cow

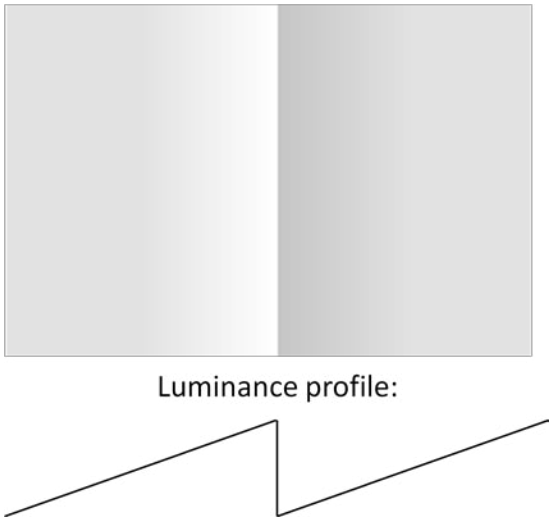

Figure 1.2: Craik-O'Brien-Cornsweet illusion

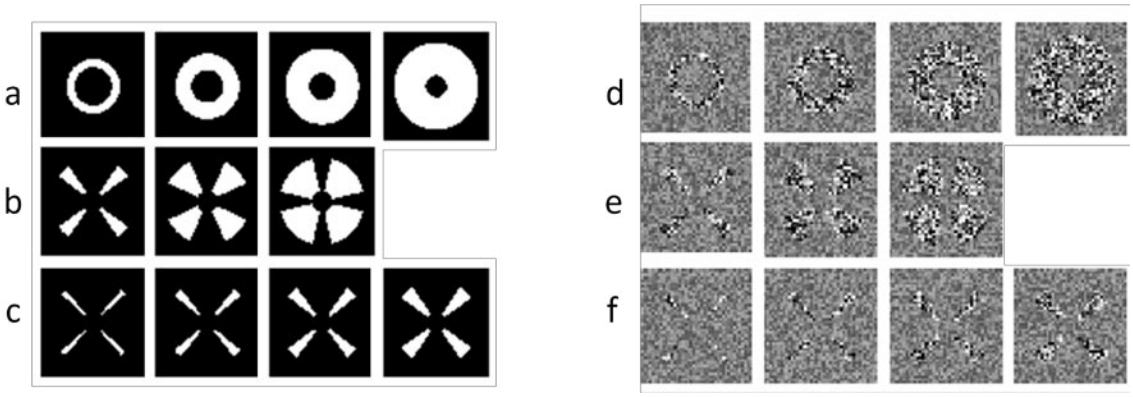

Figure 1.3: Stimuli: Left (a-c) shows just the shape template, right (d-f) shows the actual stimuli, the product of white noise at pedestal contrast (0.3) in the background, and incremented contrast inside the shape (0.7). (a,d) show 4 sizes of annuli. (b,e) show the propellers equated to the first three annulus areas. (c,f) show the propellers equated to the annulus perimeters. The propeller internal and external radii match the largest annuli radii.

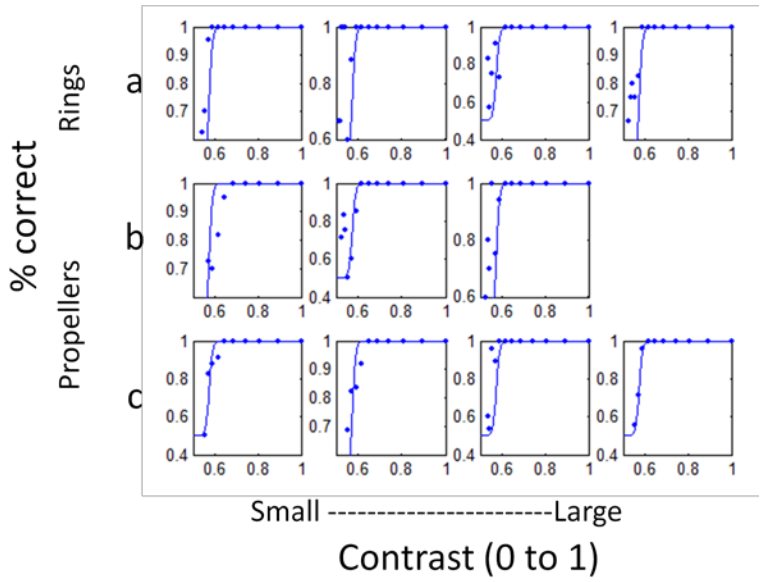

Figure 1.4: Fit psychometric curves for one sample subject. Row (a): ring stimuli. Row (b): area-matched propeller stimuli. Row (c): perimeter-matched propeller stimuli. Data are shown for each stimulus size from smallest (left) to largest (right) for each shape. Dots indicate actual data. Lines indicate fitted psychometric function. X-axis delineates the target contrast, and y-axis delineates performance as % correct. Threshold was estimated at the 71% correct level from the fitted psychometric curve. Fits were performed on individual data sets, so that there are 3-6 threshold estimates for each subject.

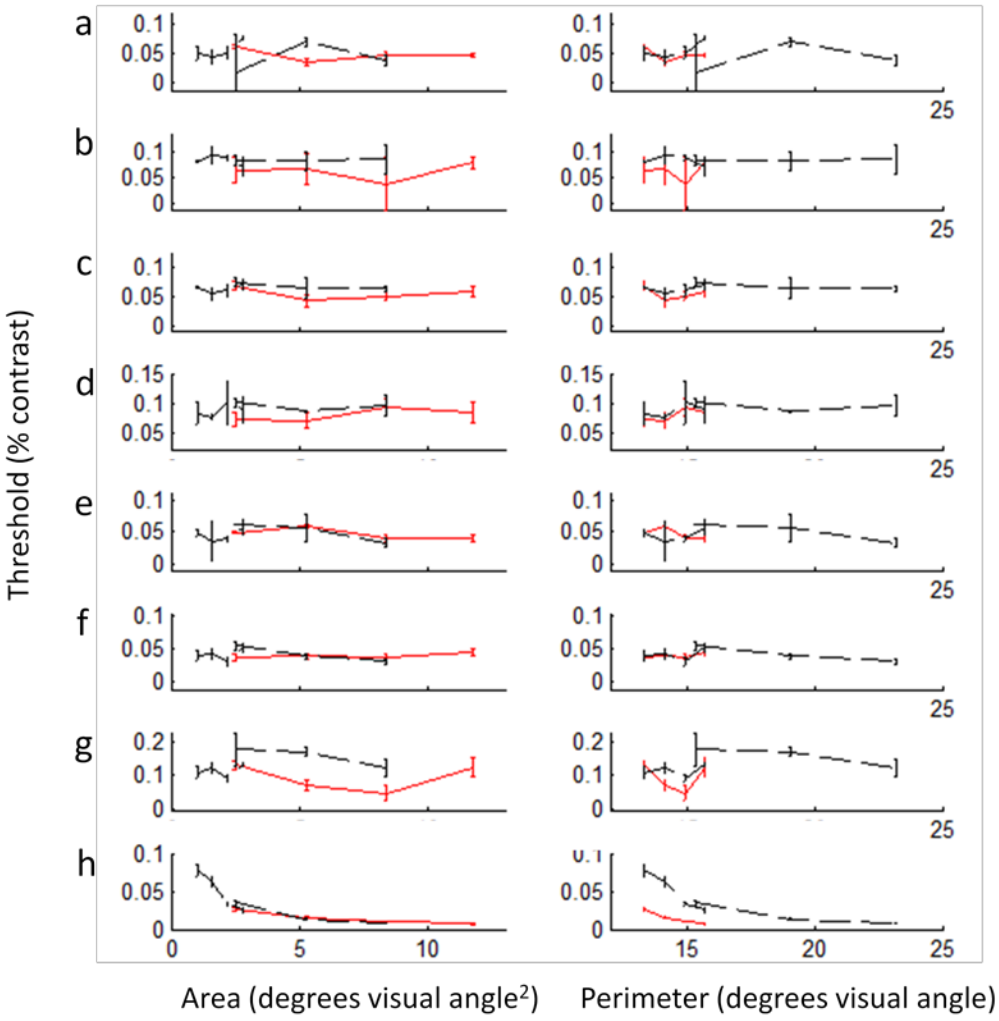

Figure 1.5: Threshold versus area (column 1) and threshold versus perimeter (column 2). (a-g) Individual subject detection thresholds. Red solid = ring stimuli, Black dashed = propeller stimuli. (h) shows the ideal observer response. Error bars are calculated from 3 separate threshold estimations and fits for subjects 1, 2, 4, 5, and 7. Six sets of data were collected for subjects (c) and (f). Error bars were calculated from 10 simulations for the ideal observer. Error bars denote standard error of the mean for each subject.

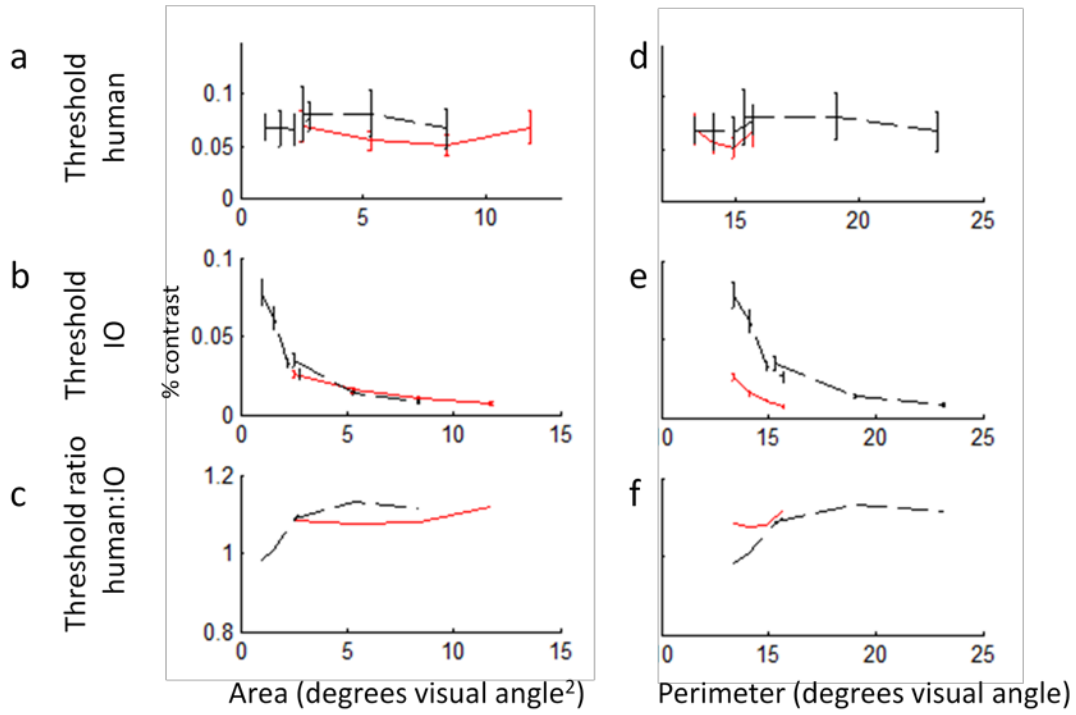

Figure 1.6: Mean thresholds versus area (column 1) and versus perimeter (column 2): human subjects (a, d), ideal observer (b,e), and ratio (c,f). Red solid = ring stimuli, Black dashed = propeller stimuli.

Chapter 2: Second-order contrast BOLD response in V1

Although the blood oxygenation-response dependent (BOLD) fMRI response in V1 has been characterized extensively for responses to a variety of both low-level and high-level visual features such as luminance, contrast, orientation and object sensitivity, relatively few studies have measured V1 response to second-order statistical properties. Because edges are rich sources of information about the substance of our visual experience, we would like to better characterize the V1 response to a variety of edges. We used BOLD fMRI to measure V1 modulation to the outside, edge and inside regions of a disk defined by either first-order (luminance) or second-order (variance) contrast. Images presented to the observers in the scanner were composed of dynamic white noise, and disks were presented in a block-alternation paradigm by modulating either the mean or the variance of the noise outside a central disk subtending 4 degrees of visual angle. For first-order disks, pixel luminance values in the background region were sampled from a distribution with a decreased mean during “on” blocks; for second-order disks, the mean was consistent throughout the entire image and pixel luminance values in the background region were sampled from a distribution with a decreased variance during “on” blocks. With this design, the mean and variance of the central disk region never changed. We measured a strong BOLD modulation in the background image regions when variance was decreased during “on” blocks, but no modulation in response to the luminance decrease in the first-order contrast condition. We found a transient response at the disk edge to both the first-order contrast-defined disk onset and offset that was

stronger in V2 but present also in V1. We found no enhancement of the BOLD response in the ROI corresponding to the disk interior for either condition.

1. Introduction

In order for V1 responses to effectively encode boundary regions in natural images, V1 should be sensitive and differentially responsive to edges created by a variety of contrasts; for example, second-order luminance contrast or orientation contrast or complex texture contrasts. We have two questions: 1) how does V1 respond to luminance- vs. contrast-defined edges and 2) is the response to texture inside the figure modulated?

Human observers are differentially sensitive to first- and second-order contrast

Although changes in average pixel luminance (first-order contrast) often provide a good estimate of important edges within an image, there is some evidence that second-order information (variance) may also provide important border information. Shofield (Schofield 2000) has found evidence suggesting that second-order information is statistically independent from first-order information. On the other hand, Lindgren et al (Lindgren et al. 2008) have argued that first- and second-order image statistics are not independent. The human visual system may benefit from an independent analysis of this second-order information, and there are a variety of psychophysical, electrophysiological, and imaging studies suggesting that first- and second-order luminance contrast are processed independently in visual cortex, or using a two-channel model. Luminance- and contrast-defined stimuli have different spatial summation characteristics: compared to

luminance-defined blobs in static white noise, contrast-defined blobs were found in psychophysics experiments to have larger spatial summation areas at both the fovea and in the periphery up to 10 degrees eccentricity (Sukumar 2007). Schofield and Georgeson (Schofield and Georgeson 1999) also found evidence that a two-channel model matches human performance better than a one-channel model in the task of detecting white noise modulated within a sine wave carrier by either pixel luminance, variance, or both.

There are a variety of other visual domains in which one- and two-channel models have also been suggested to explain discrepant results of luminance- and contrast-defined stimuli. Psychophysical studies in steroacuity suggest that a single-channel model does not explain differences observed in depth response or response susceptibility to noise when depth is defined either by luminance or contrast (Wilcox and Hess 1995; Wilcox and Hess 1997; Wilcox and Hess 1998). Motion perception studies propose that first- and second-order defined motion cues are also processed by different mechanisms: while thresholds for direction and orientation in either static or dynamic noise are the same, direction thresholds are 50% higher than orientation thresholds when motion is defined by second-order contrast (Smith and Ledgeway 1997). We predict that our results will also support the existence of a two-channel model in which first- and second-order contrast are processed independently.

Additionally, first- and second-order contrast detection thresholds themselves provide evidence for independent processing. A dipper function similar to the Legge and Foley (Legge and Foley 1980) description of the relationship between pedestal luminance contrast and contrast increment thresholds describes the relationship for detection

thresholds of luminance- and contrast-defined sine waves against increasing pedestal luminance and contrast (respectively); however, increasing luminance pedestals do not affect contrast detection thresholds, nor do contrast pedestals differentially affect luminance detection thresholds (Schofield and Georgeson 1999), again suggesting different cortical mechanisms for processing luminance contrast and second-order contrast statistics.

Finally, from magnetic and visually evoked potential studies, we know that texture-defined checkerboards elicit responses at different places over the skull than luminance-defined checkerboard elicited responses (Srebro and Baitch 1991), suggesting the importance of a more localized probe of early visual cortical response to first- and second-order contrast.

Evidence for V1 sensitivity to 2nd order contrast

While LGN cells have receptive fields much smaller than the spatial extent of texture elements and are not themselves sensitive to global texture differences, cat LGN cells have been shown to respond differentially to identical luminance patterns embedded in different types of textures and are especially sensitive to differences in spatial frequency and line terminators, suggesting that LGN does relay the appropriate information for second-order texture processing to occur (Nothdurft 1990) and that V1 may be a logical candidate for texture processing.

V1 has the processing components for texture (second-order or higher) edge selectivity. There is evidence that 2D Gabor filters (which approximate the receptive field of simple cells in the striate), can both detect different textons and find texture borders

(Turner 1986), suggesting that early visual cortex is similarly capable of texture edge detection. Using visually evoked potentials (VEP), Purpura, Victor and Katz have shown that there are numerous V1 neurons sensitive to fourth-order statistics (an edge created by adjoined even and odd isodipole textures), suggesting that V1 neurons may form the basis for higher-order texture edge detection (Purpura et al. 1994). Finally, Larsson, Landy, and Heeger used fMRI to show that there are separate populations of neurons in V1 that adapt to first and second-order contrast (Larsson et al. 2006), predicting that V1 neurons process first- and second-order statistics independently.

Neurons in V1 tend to evoke larger responses to luminance-defined edges than to edges defined by second-order or higher-order contrasts. Humans detect differences between luminance-defined edges approximately four times better than differences between texture-defined (fourth-order) edges (Victor et al. 2005). Luminance-defined edges also elicit a greater figure/ground interaction than do second-order responses in early visual cortex (Appelbaum et al. 2008; Vildavski, and Norcia 2008). More interesting, while texture (second-order) magnetic encephalographic (MEG) responses have been found to remain constant over stimulus size (when using a checkerboard), luminance defined checkerboard MEG responses were found to be greater for 0.25° stimuli than for 1° (Press et al. 2001 Wade; Wandell 2001). This neural response dependence on contrasts defined by first-order as well as second-order features suggest that V1 is processing both first- and second-order statistical properties.

To quantify the local V1 response to edges defined by first- and second-order contrast, we measured the V1 and V2 response to a target disk when defined by

luminance and when defined by contrast. Although there are many textures over which we can equalize luminance and contrast, we opted to embed our target disk in white noise because it is easy to define mathematically, easy to generate, it allows us to use novel stimuli (if scrutinized pixel by pixel) on each presentation, and yet it provides a rich percept of uniformity and is easy to segment when borders are created by regionally changing the mean pixel luminance or variance statistics.

V1 responses to contours and figure/ground segmentation

V1 shows selective activity to elements that belong to contours (which, arguably, comprise edges). Li, Piech and Gilbert (Li et al. 2006) showed that monkeys perform better on a contour detection task, and V1 neuronal response was higher, when line segments within a particular neuron's receptive field were part of a contour compared to when the elements were in the same location and orientation, but not part of a contour. Kourtzi and Huberle (Kourtzi and Huberle 2005) embedded closed contours of Gabors in distracter Gabors, and found that the fMRI BOLD adaptation to global contour structure (shape) did not cause adaptation to local element orientation (with shape constant), suggesting that V1 is sensitive to contours or shape. The corresponding evoked response to the stimuli containing embedded global contours in Gabor fields was more localized and transient in V1 than in higher areas (LOC). They postulated that a contour-dependent V1 response is likely mediated through a top-down mechanism because it disappears with anesthesia, requires training, and is retinotopically specific (Li et al. 2008). V1 sensitivity to shape may also affect cortical modulation at areas distant from the site of an edge. A few examples are 'filling-in,' figure/ground sensitivity, and texture edge response.

The mechanism of ‘filling-in’ proposes that when part of a shared surface, the neural response to edges can affect the response at the shared surface interior, as well. Lightness studies show that neural response (V1 single unit recording in Macaque and cats) starts at the edges and affects the perceived lightness interior to the edge, and response summation of multiple stimuli cannot be predicted by a linear summation of responses to individual stimuli (Paradiso et al. 2006). The phenomenon of filling-in is supported by the temporal dynamics of the psychophysical response to a luminance disk. There has been evidence in contradiction to the existence of the filling-in phenomenon. Cornelissen et al found no evidence of filling-in after measuring the fMRI BOLD signal to a block design experiment containing a disk made apparent through illusory flicker as the background was modulated between light and dark luminance levels (Cornelissen et al. 2006). However, subsequent BOLD fMRI studies have been consistent with a brightness-induced V1 signal, showing modulation to regions of the disk that were not actually modulated, but appeared dark because of the presence of a dark surround (Pereverzeva and Murray 2008).

Figure/ground studies, like filling-in studies, suggest that the border information informs cortical response in areas distant from the edge. Electrophysiology studies of surround modulation, for example Zipser, Lamme, and Schiller (Zipser et al. 1996) have shown that figure/ground cues change single-cell activity in V1 neurons. The modulation is strongest near the figure edge and decreases in amplitude toward the center of the figure. This response may occur at the frequency of the stimulus or be frequency doubled

(i.e., transient response to both onset and offset), as there have been reports of frequency-doubled edge selective VEP response to changes in edge polarity (Burr et al. 1992).

Finally, texture edges provide a paradigm in which both the local and the global contrasts (or edges) are composed of identical local cues. In this way, we can observe the effects of identifiable edges, on both sides of which are identical regions. The percept of the image depends on whether or not texture edges exist; the neural response to the image also depends on whether or not the textures edges exist. Texture edge studies also show that an early VEP component due to the onset of newly segmentable edges occurs earlier than when the edges have been maintained but the local orientation of all of the texture elements is rotated by 90° (Romani et al. 1999 and Schintone 2000). The authors interpret this latency difference (80-120 msec), which occurs with the same approximate latency as that expected due to neuronal firing (30-40 msec) (Lamme 1995), as evidence of an early edge-sensitive response informing the later sustained response to the full field.

The present study quantifies the V1 response to both the edge and interior of figure regions for disks defined by either first- or second-order luminance contrast. In addition to quantifying the relative magnitudes of the V1 responses to first- and second-order contrast when equated for perceptual salience, we tested the hypothesis that the magnitude of modulation in the interior and exterior figure regions was independent of the magnitude of modulation at the boundary region.

2. Methods

Participants

Four experienced psychophysical observers (3 female, 1 male) participated in this study. One participant was naïve to the experimental aims, while the other three were authors. All participants had normal or corrected-to-normal visual acuity. Participation was voluntary and subjects provided written, informed consent prior to participating; all procedures were in compliance with the University of Minnesota Institutional Review Board guidelines

Visual apparatus and stimuli

Visual stimuli were created using Matlab and the Psychtoolbox. Stimuli were presented on a NEC 2180UXi monitor with a resolution of 1024 x 768 pixels and a refresh rate of 60 Hz. The luminance of the monitor was calibrated using a Photo Research PR-655 photometer. The monitor had luminance range from 0.3-214 cd/m^2 with a mean luminance of 107 cd/m^2. Subjects viewed the stimuli via a mirror mounted on the head coil in the scanner.

A target disk that subtended 4° visual angle was displayed at the center of the screen. Visual noise was presented over a square extending the height of the monitor (approximately 9° height and width), with each pixel independently drawn from a uniform distribution. The distribution from which the target was drawn differed from the background distribution according to one of two conditions (luminance contrast or variance contrast, from here on out called contrast). In both conditions, there were ‘on’ and ‘off’ screens. During the ‘on’ blocks, the target could be distinguished from the background. In the ‘off’ blocks the target and background pixels were drawn from identical distributions.

Before the functional experiment, subjects completed a brief two-alternative forced choice (2AFC) task using the method of adjustment to subjectively equate the salience or visibility of targets defined by either luminance or variance. In each trial, there were two square fields of noise (each subtending three degrees) presented equidistant from the vertical center of the screen, and symmetrically with respect to the horizontal axis. In each square there was a 2 degree disk. One disk was always defined by contrast. In this square, the background had a mean luminance of 0.5 (on a scale from 0 to 1) and pixel luminance values were uniformly distributed with a variance of 0.25 (between 0 and 1). The disk had a mean luminance of 0.5 and a variance of 0.75. The other disk was defined by luminance. In this square, the background luminance was always 0.5 and the variance was always 0.25, like the background of the contrast defined disk. The locations of the luminance and contrast disks (left or right of fixation) were randomized on each trial. The mean value of the luminance disk was updated on each trial based on the response of the previous trial. If the subject reported the contrast defined disk as more salient, a 5% luminance increment was added to the luminance defined disk. If the subject reported the luminance defined disk as more salient, a 5% luminance decrement was subtracted from the luminance defined disk. Subjects were allowed to continue to adjust the luminance of the disk until they decided that they were satisfied. Each subject repeated this process four times. A point of subjective equality was thus determined and the luminance derived from this process was used in the luminance defined disk during functional scans.

In one set of scans, the target was defined by luminance (Fig. 1A). During the ‘on’ blocks, the mean luminance of the center disk was the value determined in the method of

adjustment 2AFC task described above ($\mu+\delta_\mu$), while the luminance of the background was 0.5 (μ) (identical to the background used in the above task). During the 'off' blocks, the mean luminance of the background increased to match the disk center ($\mu+\delta_\mu$), rendering the disk invisible (and effectively not present).

In the other set of scans, the target was defined by contrast (Fig. 1B). During the 'on' blocks, the variance of the center disk was 0.75 ($\sigma+\delta_\sigma$), while the variance of the background was 0.25 (σ). During the 'off' blocks, the variance of the background increased to match the disk center ($\sigma+\delta_\sigma$), rendering the disk invisible. In both types of conditions, each pixel luminance was updated at a rate of 6 Hz to help avoid retinal adaptation.

fMRI scans

For all conditions, stimuli were presented in a block design. During each scan, stimuli alternated through 10 cycles of 12 seconds 'off' and 12 seconds 'on', preceded by 12 seconds (a half-cycle) 'on' (this initial 'on' block was discarded before analysis to remove transient BOLD effects). Subjects were instructed to maintain fixation at the center of the screen. To aide attention and fixation, a rapid serial visual presentation (RSVP) task was presented at the center of the screen. A letter was presented each 200ms. Subjects were instructed to press a button on a handheld button box each time the letter was an 'x'.

In order to independently localize the region of cortex corresponding to the retinotopic location of the background, edge, and interior of the disk, we used a differential localizer (Olman, Inati and Heeger 2007). We used the same block paradigm

as above, however, while ‘on’ blocks displayed a series of concentric rings with inner and outer radii of 0-0.5°, 1-1.5°, 2.5-4° visual angle, the ‘off’ blocks displayed a complimentary series of concentric rings with inner and outer radii of 0.5-1°, 1.5-2.5°, and the entire screen outside of 4° visual angle. The rings were filled with a 100% Michelson contrast polarity-reversing checkerboard modulated at a rate of 5 Hz. Mean gray filled the spaces complementing the rings.

Data acquisition and pre-processing

All MR images were acquired with gradient echo EPI (TR 1.5s, TE 20 ms) at 3T (Siemens Trio system) using an 8-channel volume coil. Slices were oblique and parallel to the calcarine sulcus in the occipital cortex. The field of view was 256 x 128 with a matrix size of 128 x 64 voxels (outer volume suppression was used to suppress signal outside the specified field of view), yielding a resolution of 2mm x 2mm x 2mm with 18 slices total. Data were reconstructed using the native Siemens software. To compensate for distortion caused by magnetic field inhomogeneities, all functional data were unwarped using a field map that was acquired during the functional imaging scanning session.

Retinotopic mapping

Primary visual cortex (V1 and V2) were identified using standard retinotopic mapping techniques (Sereno et al. 1995; DeYoe et al. 1996; Engel et al. 1997). To map eccentricity we used dynamic checkerboard stimuli masked by expanding and contracting rings, with rings either expanding from 0° to 12° over 24 seconds or contracting from 12°

to 0° over 24 seconds. To compensate for hemodynamic delays, data from contracting-ring scans were time-reversed and averaged with expanding-ring scans (after shifting both scans forward in time by 3 TRs to partially compensate for the hemodynamic delay and increase contrast-to-noise ratio after averaging). After calculating the response phase and coherence for each voxel in the average (delay-free) scans, we used the phase values as an estimate of response time and thereby identified eccentricity representations in retinotopic areas. We used a 45° sliding checkerboard wedge that rotated about fixation (completing one cycle in 24 seconds) to elicit traveling waves of activity mapping polar angle across each of the retinotopic visual areas. After similarly calculating the phase and coherence of each voxel, because the phase (or temporal delay) corresponds to the polar angle of the stimulus wedge, we defined V1 as the region centered on the calcarine sulcus representing the contralateral hemifield, and V2 as the areas corresponding to the lower contralateral quarter field above V1 and the upper contralateral quarter field below V1.

Alignment to reference volume anatomy

After motion compensation (custom Matlab tools) and distortion compensation (FSL, http:// www.fmrib.ox.ac.uk/fsl/), the first volume of the first EPI series was then aligned to a 3D MP-RAGE volume of the whole brain collected in a separate scanning session. The voxel intensities in the EPI data were inverted to match the T_1 contrast of the anatomical data, and automatic image alignment (Nestares and Heeger 2000) was used to align EPI data with the reference anatomical data.

Segmentation of cortical surface

The gray/white matter surface was segmented and reconstructed from the MP-RAGE volume anatomy of each subject with SurfRelax (Larsson 2001). Patches of the occipital cortex were cut from the inflated whole brain and flattened to aid visualization of the functional data.

Data analysis

Data analysis was performed with code written in MATLAB. The first 8 time points (1/2 cycle) were discarded and motion compensation was performed on the remaining data from each scan (Nestares and Heeger 2000). High pass filtering was then performed on the data with a cut-off frequency of 4 cycles per scan.

Repeated scans of the differential localizer were averaged together and regions of interest (ROIs) were defined to include voxels coherence (unsigned correlation with a sinusoid at the stimulus alternation frequency) greater than 0.3 during the differential localizers. We then confined ROIs defined by localizers to V1 using the previously collected retinotopic mapping data aligned to the same reference anatomy.

3. Results

BOLD fMRI peak luminance and contrast response patterns

Each subject completed 2 or 3 scans with 10 cycles per scan, or between 20 and 30 repetitions, for each condition. There were three conditions: the localizer (rings vs. complementary rings in visual space), luminance-defined disks, and contrast-defined disks. For each subject, the localizer produced BOLD modulation in the early visual cortex that was in phase with either the set of annuli presented in the first half of each

block or in phase with the annuli presented in the second half of each block. This is seen in figure 2.2c as either orange regions (in phase with presentation of one set of rings) or blue regions (in phase with presentation of the other set of rings), visualized on the flattened cortical surface. Voxels were assigned to the edge ROI if they fell into the second orange band of activation in V1, which corresponded to the annulus at 1.5 to 2.5° eccentricity. Voxels were assigned to the interior disk ROI if they fell into the first three bands of activation (blue, orange, blue), and into the disk exterior ROI if the fell into the third blue band or the third orange band. Band activation was continuous in V1, V2, and V3. Because the hemispheres, although not completely independent measurements, are spatially independent responses to the stimuli, we assigned separate ROIs to the right and left hemispheres and analyzed them as independent response measurements.

To characterize the magnitude of the BOLD modulation, we used the peak amplitude of the mean response averaged through all blocks in one scan, and in all scans per subject. Mean hemodynamic response functions (across subjects) are shown in figure 2.3 a and b at the stimulus frequency and 2.3 c and d at the first harmonic (sensitive to transient responses). Because there was previous work showing a heterogeneity in first-order and second-order sensitive neural populations (Larsson et al. 2006), which may not peak at the same time, our analysis based on peak response may be more biased towards the stronger responding population. Therefore, we also completed analysis of the individual voxel Fourier components at both the stimulus frequency (10 cycles) and the first harmonic (20 cycles). We used the average Fourier components through the voxels within each ROI as our measure of stimulus modulation.

Because we are estimating BOLD response as the magnitude of the Fourier component at the stimulus-alternation frequency, the null hypothesis for statistical tests for significant modulation is not a value of zero, but a positive value representing the magnitude of the stimulus-related Fourier component in the noise spectrum. We estimated the expected modulation due to noise by first converting the ROI data to the frequency spectrum, then removing the components at the fundamental and first harmonic of the stimulus presentation frequency, fitting the spectrum with a third order polynomial, and using the fitted (baseline) value at the fundamental and first harmonic as the expected noise to both the stimulus, as well as the onset/offset of the stimulus. Our first analysis was to determine which ROIs contained modulations larger than noise.

ROI-based analysis: Linearity of response to contrast-defined stimuli

Contrast-defined stimuli resulted in significant BOLD modulation in both the edge and the outside ROIs (figure 2.5a, middle column; 2.5c, middle panel). This observed response is expected from first-order contrast sensitivity in V1. The edge response was very close to a linear combination of the inside and outside ROI responses. Figure 2.6 shows both the linear regression for each individual subject, as well as the linear regression of the mean through subjects in the frequency domain. Both pairs of regressions show highly linear relationships between the ROI responses.

ROI-based analysis: Luminance-defined stimulus response

Luminance (whether inside or outside the figure) very weakly modulated the BOLD response as compared to the contrast first-order contrast, consistent with previous

findings (figure 2.5a, left column; figure 2.5c, left panel). However, the luminance-defined contrast (first-order contrast) did elicit a frequency doubled modulation at the disk edge (figure 2.5 b and d, left panels). Based both on this observation, and on the lack of first harmonic response in the outside region in the contrast modulated condition (figure 2.5d, middle panel – this region represents a first-order contrast manipulation in the contrast-defined stimulus condition), we may conclude that the figure elicits a unique, first harmonic sensitive first-order contrast response.

Of note, there was a significant difference between the groups of modulations measured across the 'inside' ROI and the modulations estimated to occur due to random noise. However, both the mean response as well as each individual subject measurement was smaller than the mean noise estimate and each individual noise estimate, respectively. The difference between the means was much smaller than one standard error of the mean, and probably represents a conservative estimate of the noise or a poor fit to the estimated noise as a function of frequency component. A few contributing factors to this overestimation may be the inclusion of higher harmonics in the fitted noise spectrum, or an inappropriate use of a third-order polynomial for the function used to fit the spectrum.

Hartley's statistical analysis

Because our first analysis relied on an estimate of the spatial correlation between voxels, which was an approximation, we also tested for significant modulation using Hartley's test, or the F_{max} test. This test compares the ratio of the relative power in the data at the signal frequency to the average power over the frequency spectrum omitting

the frequency component of interest (Hartley 1949, 1950). The test assumes the data are drawn from a normal distribution with equal members in each group. Using this statistic, we find additional signals significantly above the estimated noise. The luminance condition yields signals in the edge ROI at both F0 and F1, and the outside at F0 ($p<0.05$). The contrast condition yields signals at the disk inside and outside at F0, at the disk edge for both F0 and F1 ($p<0.05$). Because the power spectrum of our data is not flat, the Hartley test may provide an overconfident rejection of the null hypothesis (that signal at the frequency of the stimulus is not different from noise).

BOLD fMRI peak luminance and contrast response over cortical distance

To further investigate the fundamental response of the cortex to a second-order edge without blurring across an entire region of interest, we mapped the contrast response as a function of cortical distance from the fovea (figure 2.3). In both the luminance and contrast conditions, we saw very little evidence of vascular blurring inside of the figure. As illustrated in the above figure (2.3a and 2.3c), there is a strong response that begins at the disk edge (indicated by the white band in 2.3b), and rapidly increases in the contrast condition (far right), with strongest modulation far into the disk exterior. The response to second-order contrast at the disk edge as quantified by our ROI analysis could be due to blurring across a coarsely defined area, and this blurred signal could overpower a more subtle response to second-order contrast that may be present in a subset of more retinotopically selective voxels.

4. Discussion

Expected response to luminance

It has been shown that V1 and other early cortical areas respond to surface luminance changes (Haynes et al. 2004). However, we did not observe any luminance response modulation. There may be several reasons that explain our observations that are consistent with what is already known about V1 response to luminance. Haynes et al. measured V1 response to luminance changes was positive for both increments and decrements in luminance, and larger for decrements than for increments. We did not observe a similar transient response to luminance change in our outside ROI. Because we added dynamic noise, we may have abolished any luminance response or sustained the response throughout the experiment. Secondly, Haynes et al used a static, uniform luminance with a minimum luminance difference of 30 cd/m^2 and measured a transient response that decayed rapidly over their 15 second test. Our average luminance change (5 cd/m^2) was also much smaller than the luminance modulation because we matched our luminance values to contrast defined disk saliency. Finally, there is recent evidence that the V1 response to luminance corresponds to lightness perception. Although there was a brightness change in the surround, the perceived lightness may have been diminished by the disappearance of the disk concomitant with the luminance change or masked by noise (Boyaci et al. 2007; Pereverzeva and Murray 2008).

Figure/ground modulation and second-order contrast response

Neural response can change when the neuron's corresponding receptive field is a part of a figure or a figure edge. These phenomena include border ownership sensitivity and figure/ground modulation. There is some evidence that V1 neurons exhibit border

ownership sensitivity. Zhou, Friedman, and von der Heydt (von der Heydt et al. 2000) found a small number of V1 neurons that exhibit feed-forward mediated border ownership sensitivity. The same group, however, suggest that figure/ground sensitivity does not start until V2 (von der Heydt et al. 2000). Figure effects are also evident in studies that modulate a target figure's local context (or border contrast) and find that V1 neurons corresponding retinotopically to the figure interior change as a function of context. These figure/ground studies span a variety of methodologies: electrophysiology texture studies (Lamme 1995), evoked potential to texture (Lamme et al. 1993 Spekreijse 1993), evoked potential to coherent motion-defined detectable contours (Lamme et al. 1993), and human fMRI, (Skiera et al. 2000 and Fahle 2000; Fang et al. 2009).

Our results are consistent with a model in which V1 response to both first- and second-order contrast is sensitive to onset and offset, possibly due to the appearance and disappearance of a figure. An alternative explanation is that a second-order contrast-sensitive response exists that is overwhelmed in the presence of first-order contrast, but occurs at the onset and offset of the second-order contrast stimulus. Given that our data are consistent with the edge response to contrast-defined stimuli being the mean between the responses inside and outside the disk, this first-order contrast response dominates the response at the stimulus frequency. However, the first harmonic response at the edge is stronger than the first harmonic response in either the inside or the outside regions, suggesting that the second-order contrast response is independent of the first-order contrast response, transient, and sensitive to both the figure onset and offset.

Although we did not see any positive figural effect represented in the area of V1 corresponding retinotopically to the disk interior, our results do not contradict the evidence that V1 shows figure/ground sensitivity. One possibility is that the V1 response to the dynamic noise was strong enough to mask a weak modulation based on figure-ground segmentation. A second consideration may be that the signal was strong, but the after effect reported by all subjects was strong enough to equal the figure perception and make the figure and after image figure responses indistinguishable in the BOLD response. Another possibility is that we are seeing reduced activity due to the presence of a discernable figure, similar to results found by Murray et al (Murray et al. 2002).

BOLD fMRI peak luminance and contrast response over cortical distance

A transient response could be due to an averaging across the response to the disk exterior and the disk interior. This would cause the peak (and pre-adaptation) response in the disk exterior to appear to look like a transient response in the disk edge at the disk appearance. In both the luminance and contrast conditions, we saw very little evidence (according to the response as a function of cortical distance in figure 2.4) of vascular blurring corrupting our results. Because the penetrating intracortical venuoles are spaced at approximately 20 um for the smallest of five venule sizes to 0.5-2 mm for the largest of the venule sizes, which is roughly on the same order as our voxels (Duvernoy et al. 1981), we are therefore confident that as least some of the transient response observed for second-order defined stimuli is a representation of the underlying neural response.

5. Conclusion

We find that the first-order contrast response dominates any second-order contrast response at the frequency of the stimulus presentation. Secondly, we measured a transient onset and offset response to the luminance-defined stimulus due to first-order contrast. Finally, we measured a first harmonic response to the contrast-defined stimulus that may be due either to a figure sensitive first-order mechanism independent and in addition to the main first-order response (seen at the stimulus frequency) or due to a second-order contrast mechanism.

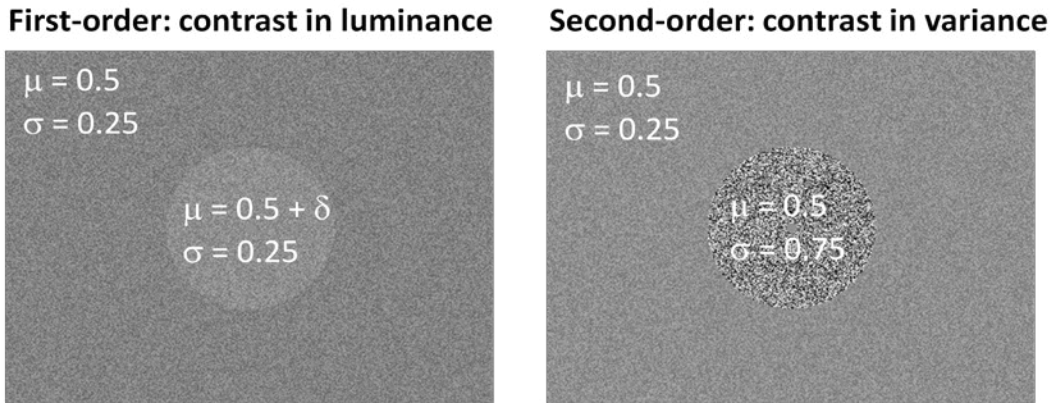

Figure 2.1: Stimuli for 2nd-order contrast fMRI experiment. Stimuli were defined a) by mean pixel luminance change, and b) by pixel variance change. Target was equalized in salience for each observer prior to the onset of the scanning session by performing a method of adjustment task to determine the point of equality for luminance and second-order contrast defined target disks. Blocks alternated every 12 seconds, for a total of 10.5 cycles.

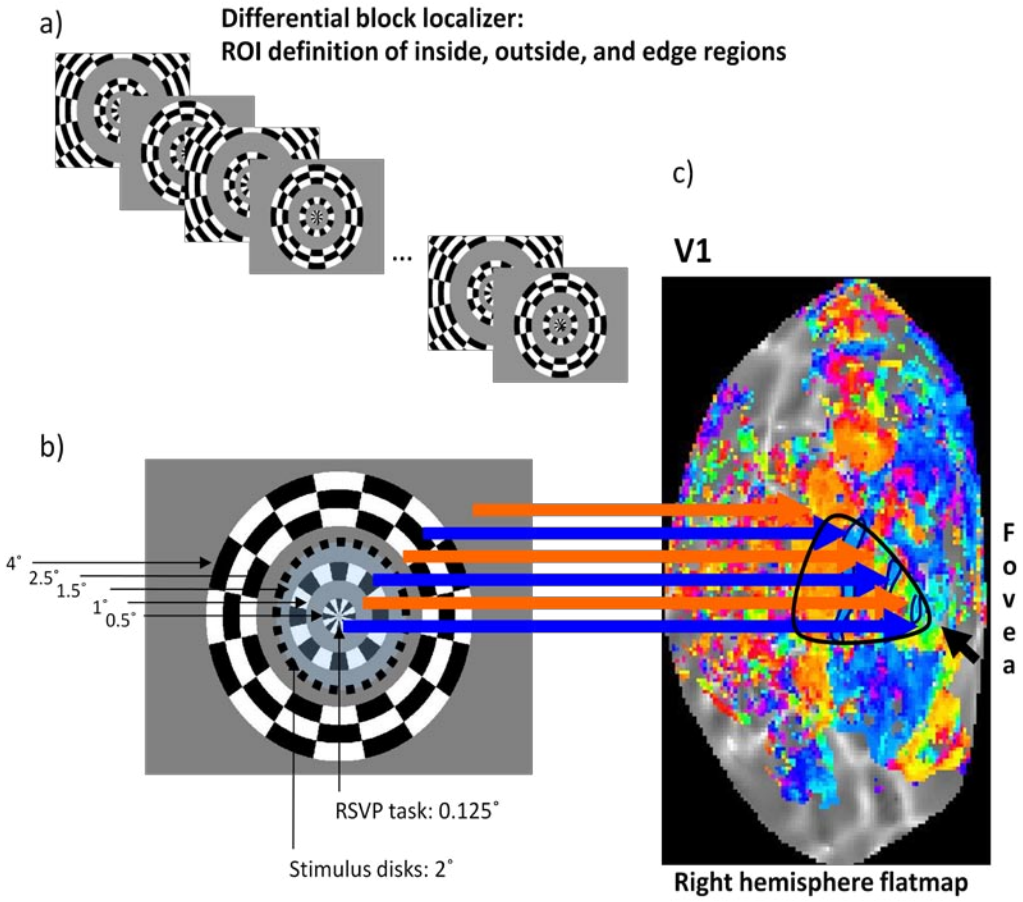

Figure 2.2: Localization of regions of interest for analysis of BOLD fMRI data. A) Differential localizer fMRI block design. This design was used to independently localize the disk interior region, the disk edge region, and the disk exterior region. Blocks alternated every 12 seconds for 10.5 cycles total. The checkerboard was displayed at 100% of the total available screen contrast, and the polarity reversed at 5 Hz. B) The localizer disk contained rings at 0.5, 1, 1.5, 2.5, and 4° eccentricity. C) A patch from one representative subject's occipital pole, inflated and flattened for viewing purposes. V1 boundaries and fovea are indicated by a closed contour and arrow, respectively. Lines separate the responses corresponding to changes in phase, which are consistent with the ring edges.

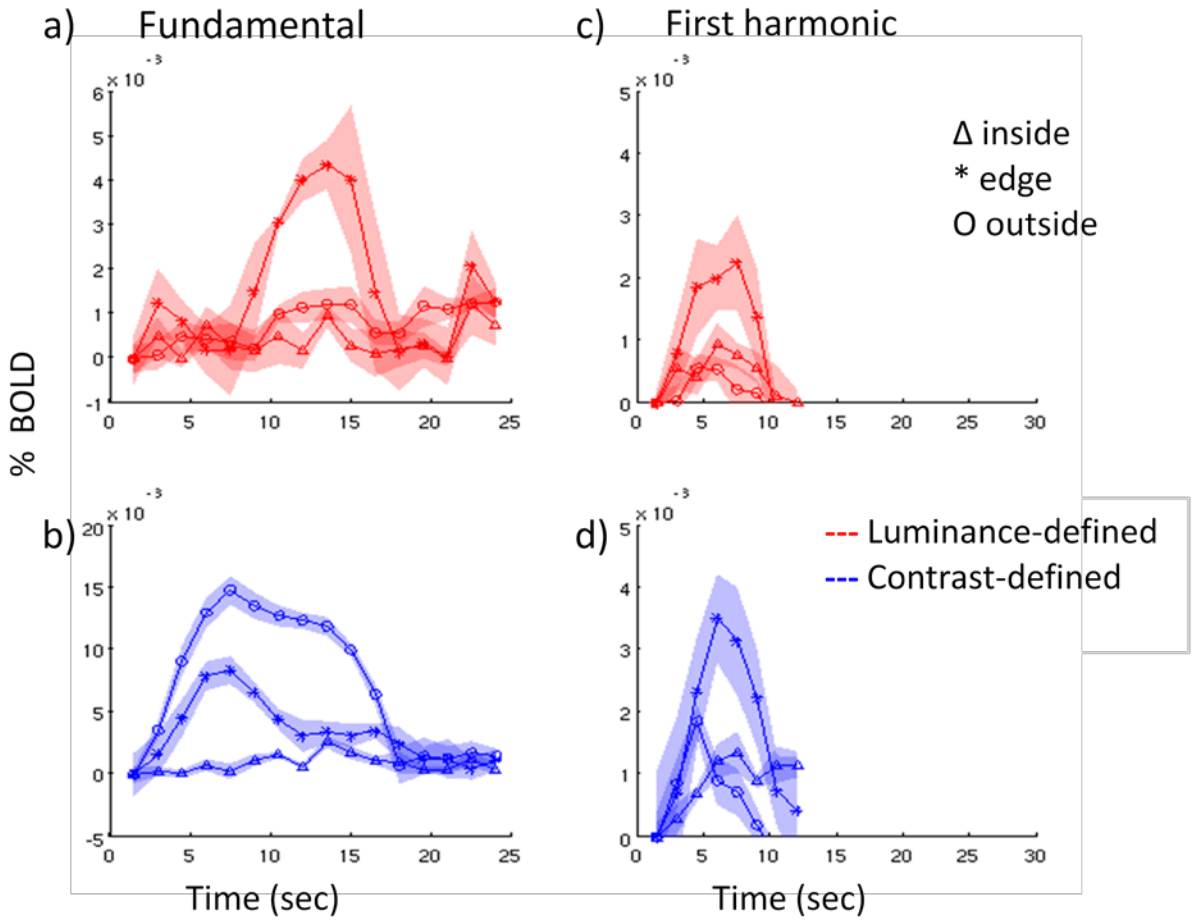

Figure 2.3: Hemodynamic responses. HRFs are estimated by averaging BOLD data in each ROI for either 24 s after stimulus onset (fundamental stimulus frequency) or for 12 s after both onset and offset (first harmonic). HRFs for the luminance condition (a and c; red) and contrast condition (b and d; blue). Each ROI average response is shown to the fundamental stimulus frequency (a and b) and the first harmonic (c and d). Shaded regions indicate SEM.

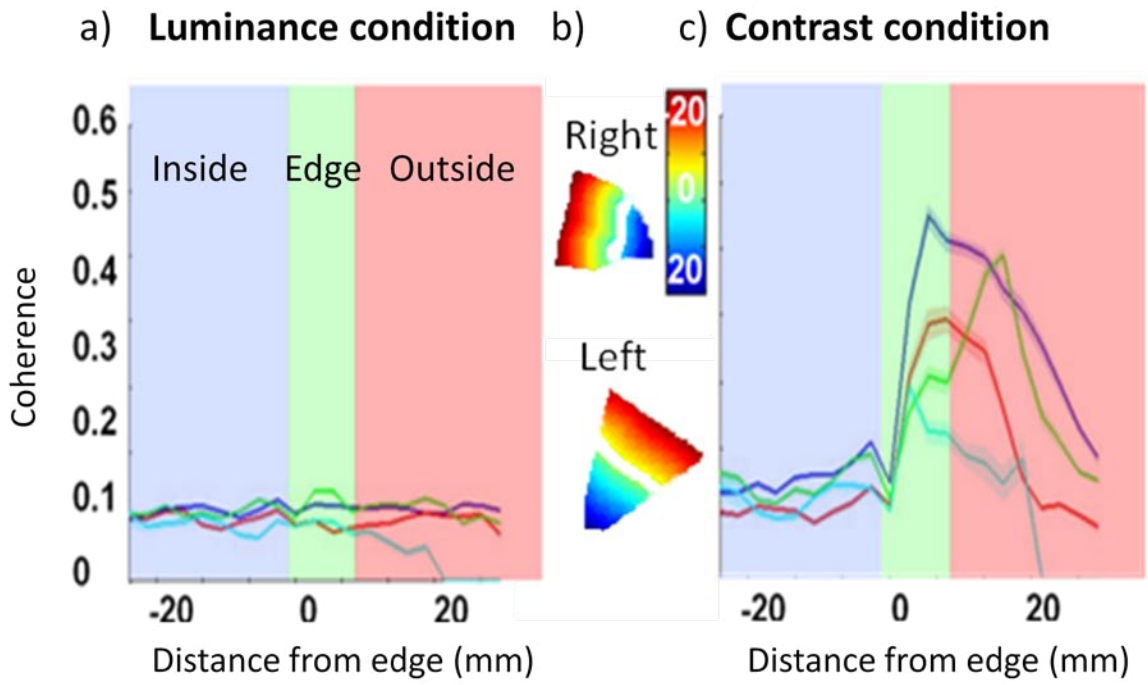

Figure 2.4: Cortical distance plots. Average response as a function of cortical distance from the edge of the disk. The x axis indicates distance from the disk edge in millimeters, while the y axis indicates the coherence. Shading indicates the ROI these voxel responses would have been binned and averaged into in figures 2.4 and 2.6 (blue = inside, green = edge, red = outside). This shading is included here for reference only. The ROIs were not used for any part of this analysis. The left panel (a) corresponds to responses to the luminance defined disk. The right panel (c) corresponds to responses to the second-order contrast defined disk. The middle panel (b) shows the inflated, flattened cortical representation of V1 for both hemispheres for one representative subject. The disk edge is shown on both hemispheres as a white line. This line corresponds to 0 mm in the left and right panels. The color overlay corresponds to the distance from the border in mm (as indicated in the phase map on the upper middle panel).

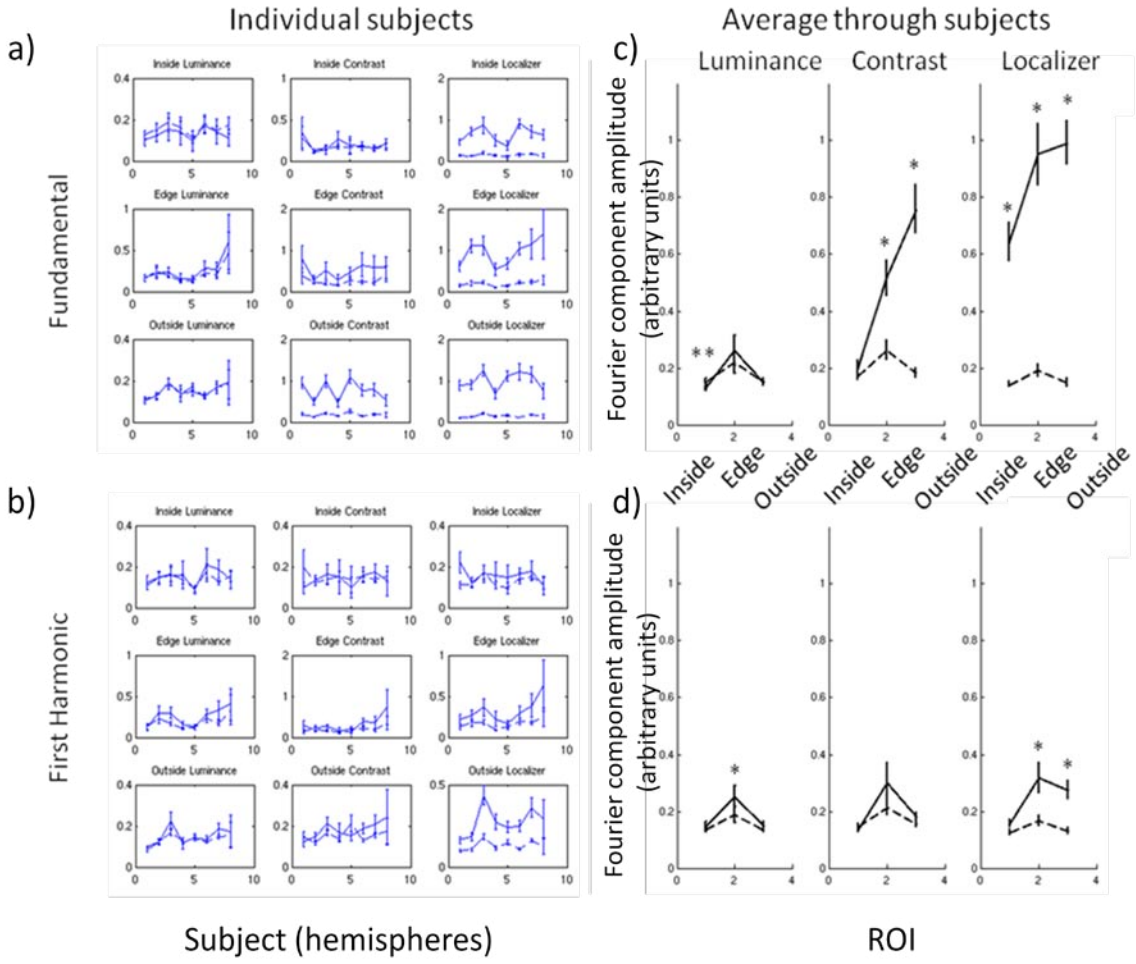

Figure 2.5: Fourier analysis. Individual Fourier component amplitude (solid line) and noise estimate (dashed line). (a and c) show the fundamental frequency component response, while (b and d) show the response at the first harmonic. For both (a) and (b), which show signal and noise amplitude estimates for 8 hemispheres in 4 subjects, column represents condition (left to right = luminance, contrast, localizer) while row represents ROI (top to bottom = inside, edge, outside). For (c) and (d), the each panel shows the ROI response (shown on the x-axis) corresponding to a different condition (left to right = luminance, contrast, localizer). * indicates a group difference across subjects from the noise estimates for each subject. ** indicates a group difference with the noise mean higher than the signal mean.

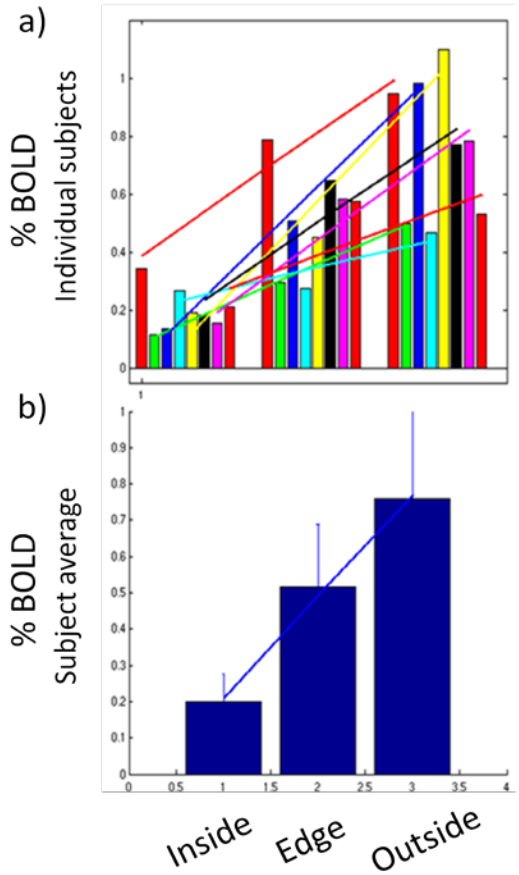

Figure 2.6: ROI comparison. Linear regression over contrast-defined stimulus modulation at the stimulus frequency. (a) shows individual subject data (4 subjects, 2 hemispheres per subject). Each color represents one hemisphere (b) shows mean response, error bars indicate SEM, solid line indicates linear regression of the mean responses.

Table 2.1: Comparison of BOLD modulation to noise estimate

	Fundamental frequency		First harmonic	
	Luminance	Contrast	Luminance	Contrast
Inside	13.084* (0.0066)	20.09 (0.1065)	15.05 (0.1913)	13.76 (0.4917)
Edge	26.17* (0.0042)	51.61* (0.0011)	25.25* (0.01097)	30.04 (0.1717)
Outside	14.84 (0.4357)	76.03* (0.00011)	14.82 (0.1523)	18.00 (0.1720)

Values for the mean frequency component are shown for both the fundamental (left two columns) and first harmonic (right two columns) for each test condition (luminance and contrast). * indicate group statistical significance above noise. The t-test p-value is show in parentheses.

Chapter 3: Temporal onset nonlinearities in the BOLD response in V1.

The timing of the hemodynamic response function (HRF) is assumed to be a linear function of the underlying neural response. As stimulus contrast increases, both retinal (Shapley and Victor 1978) and cortical (Dean and Tolhurst 1986; Albrecht 1995) neuron response latencies decrease on the order of a few tens of milliseconds, and as the contrast increases, latency also decreases (Reich et al. 2001). If the functional magnetic resonance (fMRI) measured blood-oxygenation-level-dependent (BOLD) signal is a linear transform of the neural signal, we would expect the BOLD signal contrast-dependent latency to be on the same order as this neural latency. However, there are also known nonlinearities in the BOLD signal that affect the amplitude (and possibly the timing) of the HRF that far exceed the contrast-dependent nonlinearities of the neural response. Many previous paradigms have reported that the contrast-dependent BOLD amplitude peak response violates the linear property of summation. A few studies have reported timing nonlinearities (Vazquez and Noll 1998; Miezin et al. 2000), but no paradigms have probed the timing with high temporal resolution. In the present study, we sampled the HRF elicited by 5, 10, 25, and 90% contrast, contrast-reversing checkerboards and sampled the BOLD response using a repetition time (TR) of 250 ms to test the hypothesis that small timing nonlinearities exist as a function of stimulus strength. We found that the onset latency of the BOLD response to high-contrast stimuli was approximately 300-400 ms less than the onset latency of the BOLD response to low-contrast stimuli at 3T; the mean onset latency difference was 227 msec in gradient echo (GE) data acquired at 7T. We also found that full-width half-maximum (FWHM) dependence on contrast in data

that were acquired using a spaced paradigm at 3T. The difference between 5% and 90% contrast FWHM was 45 msec, smaller for low contrast stimuli than for high contrast stimuli, which suggests that the HRF shape changes with contrast, and the latency we observed is not just due to a temporal shift in the response.

1. Introduction

BOLD reflects neural activity with a nonlinear transformation

Blood oxygen-level dependent (BOLD) signals are well established to be faithful to the blood oxygenation level (Ogawa et al. 1990), and more importantly, to some aspects of the local neural activity (Buxton et al. 1998). The relationship between stimulus contrast and response is monotonic for the BOLD peak amplitude in early visual cortex (Tootell et al. 1995; Boynton et al. 1996 and Heeger 1996), and local neural activity (Boynton et al. 1996).

Many people use the assumption that the BOLD response reflects a linear and time invariant system (Boynton et al. 1996 and Heeger 1996; Dale 1997; Heckman et al. 2007) in the analysis of event-related (ER) experiments in order to draw conclusions about the character of the neural impulse response function. This assumption allows us to present stimuli more frequently than once every 16-20 seconds, the average length of the hemodynamic response function (HRF), and still maintain the ability to separate the stimulus responses using linear methods, with the end goal of producing robust estimates of stimulus response (Buckner et al. 1996 Savoy, Peterson, Raichle, and Rosen 1996) by presenting as many stimuli in as little time as possible while maximizing data acquired while a subject is readily engaged and attentive.

While many event-related experiments have demonstrated the BOLD response to be linear (obeying the properties of separability and linear summation), linear behavior is usually limited to stimuli spaced in time, or high in contrast (Vazquez and Noll 1998). Linear separability by deconvolution is observed for stimuli as close together as 4 seconds (Glover 1999). Linear summation of the peak BOLD response may depend on both the stimulus duration and the time between stimuli. Linear summation has variously been found to break down with 1 second duration (Burock et al. 1998), 2 second interstimulus interval (Dale 1997; Buckner 1998), a 3 second interstimulus interval (Boynton et al. 1996 and Heeger 1996) or as much as 4 seconds required between trials (Vazquez and Noll 1998; Liu and Gao 2000). These findings suggest that when presented more rapidly than a few seconds between stimuli, peak amplitude of BOLD modulation decreases (Birn et al. 2001).

Conclusions about nonlinearities in the BOLD peak amplitude are found extensively in the literature; however, there are only a few findings reported regarding the linearity of the BOLD response timing. Some evidence suggests that precise BOLD timing in V1 and other visual areas is reliable within a given area, and heterogeneous between areas (Birn et al. 2001). Several studies suggest that neither time-to-onset nor time-to-peak appear to be correlated with the peak amplitude of the response (Miezin et al. 2000 Ollinger, Petersen and Buckner 2000) even for stimulus paradigms that produce nonlinear peak amplitude responses (Grinband et al. 2008), and even for stimulus durations as short as 2 ms (Yesilyurt 2008). Furthermore, response onset latencies (from here on referred to as latency) in the 10's of ms are found to correspond well with task measures such as

reaction time and stimulus presentation time (Menon et al. 1998). These findings have drawn people to use fMRI to resolve localized sub-second timing distinctions between stimulus responses and to make estimates about the relative ordering of spatial processes and causality (Friston et al. 2003). There is evidence that timing differences can be extracted on the order of 200 ms (Singh 2002 Zhou 2002), and these timing differences can be used in to make conclusions about the relative processing order or information flow between the timed areas using dynamic causal modeling (Friston et al. 2002; Friston et al. 2003). Causal modeling has been used, for example, by Ravindranath, Zhang and Mujica-Parodi (Ravindranath 2007) to make statements about the way in which the amygdala, hippocampus, and prefrontal cortex are connected, and to model the connectivity in early visual cortex (Marreiros et al. 2008; Stephan et al. 2008 Daunizeau, Ouden, Breakspear, Friston 2008).

However, there is also evidence that the timing properties of the HRF are subject to nonlinearities under the same conditions that cause peak amplitude nonlinearities. Miezin et al have shown that, while stimuli as far apart as 5 seconds suffer a 17-25% peak amplitude decrease, they maintain relative stability in time to peak, but show some instability in time to onset, even when confined to one region within one subject (Miezin et al. 2000 Petersen and Buckner 2000). Even with implementation of robust experimental designs, few HRF assumptions, and robust data analysis, Vazquez et al found substantial nonlinearities in the magnitude, peak delay, and dispersion of the hemodynamic response when using an interstimulus interval (ISI) of 1 second (Vazquez and Noll 1998 and Noll 2005) and de Zwart et al used a rapid ER GE fMRI experiment

with an ISI of 1 second and a TR of 1 second to establish that substantial nonlinearities are present causing the HRF to broaden in time and dampen in amplitude (de Zwart et al. 2009 Fukunaga, Bianciardi, and Duyn 2009).

Origins of nonlinearities in the BOLD response

Specific theories about the origin of BOLD nonlinearities have been postulated implicating both neural and vascular causes. And although neural latency nonlinearities have been observed they are not likely the only contributor to BOLD latency nonlinearities. Neural response latency has been shown to increase by about 40 ms as contrast of an optimally oriented bar decreases from 78 to 5% contrast (Gawne et al. 1996), and nonlinear contrast response compression occurs within the first and last 10 ms of the neural response profile (Albrecht et al. 2002). Budd et al (Budd et al. 1998) postulated that neural populations with varying latencies within an area could produce nonlinearities in the average neural response strength (measured by EEG) in response to stimulus repetition, but only in the tens of milliseconds. Neural responses may also contribute to BOLD responses as the population response decorrelates within the first 80-100 ms of stimulus onset (Hegde and Van Essen 2007) and, for natural stimuli, a late onset inhibitory response creates a biphasic temporal nonlinearity (David et al. 2004). Another postulated neural cause for nonlinear BOLD response to stimulus strength is neural adaptation (Boynton et al. 1996; Heckman et al. 2007).

Most theories attribute BOLD nonlinearities to vascular causes. Mandeville et al suggest that BOLD amplitude nonlinearities are due to an initial rapid capillary and venous elastic response followed by a delayed venous compliance causing the post

stimulus undershoot (Mandeville et al. 1998). Nonlinear onset and offset characteristics such as the initial dip and post stimulus undershoot have been postulated to represent a transient neurovascular uncoupling and subsequent momentary nonlinearity (Buxton et al. 1998; Uludag 2008). Both BOLD response amplitude and latency may be affected by these blood flow nonlinearities. de Zwart et al showed that during a rapid ER paradigm that resulted in HRF temporal broadening amplitude dampening, there were no changes in the MEG response as a function of stimulus time separation, also suggesting a vascular cause of BOLD temporal nonlinearities (de Zwart et al. 2009).

Although neural and vascular events both contribute to the BOLD response, transient neural responses at the beginning and/or end of stimulation giving rise to phasic (nonlinear) BOLD signals cannot be differentiated from vascular uncoupling, and ultimately, without independent measures of either neural activity or blood flow, these events cannot be infallibly differentiated.

Characterizing BOLD timing

There are also many methods that were developed to improve the faithfulness of the BOLD data to neural signal, including optimization of the fMRI acquisition methods, optimization of the experimental design, assumptions governing fMRI data analysis, and fMRI acquisition methods. Improving the canonical HRF estimate, alignment of the neural and BOLD data, or using models that rely on fewer assumptions about the relationship between neural activity and BOLD response (Lindquist and Wager 2007; Lindquist et al. 2007) can improve the accuracy of the extracted fMRI data.

For example, sampling at high temporal rates increases the accuracy of the timing information in acquired data. By sampling at 250 ms and then undersampling at different rates, Dilharreguy et al estimated that a decrease in accuracy amounts to about 50 ms per second of TR increase. Even if there is no temporal cut-off, undersampling still takes disproportionate account of confounds such as large veins and the poststimulus undershoot, suggesting that any timing studies should be carried out with high sampling rates (Dilharreguy et al. 2003). Furthermore, these authors estimate that if one uses a temporal cut-off of 3-4 seconds after the HRF peak time, there can be up to a 100 ms time shift in the HRF peak time for a 1500 ms TR. Because undersampling can misrepresent the fine timing characteristics of the HRF by introducing a up to a 50 ms onset lag, over-represent the large vascular response to the stimulus, and create up to a 100 ms shift in the peak time, our purpose is to rapidly sample the HRF to obtain reliable estimates of the timing properties of the contrast evoked HRF.

The two purposes of the current study are to clearly sample the BOLD HRF at high temporal resolution, and to address the cause of any nonlinearities found in the HRF. We measured the fine timing information of the as a function of stimulus intensity using a polarity-reversing checkerboard in both a rapid ER design and a spaced ER design experiment. The rapid ER experiment allowed us to collect responses to more contrast levels while keeping the total scan time reasonable for the observer to perform (6:18 minutes with 50 repetitions of each stimulus contrast per scan). We measured responses to 2.5, 5, 25, and 90% contrast. The spaced ER experiment ensured that our timing observations were not affected by the overlapping HRFs or the deconvolution of the

HRFs to produce event-triggered average responses. Because the individual event responses are expected to be larger in the spaced ER experiment than during the rapid ER experiment (due to the amplitude decrease observed with rapid ER experiments), we may also measure better timing estimates because the expected evoked response is larger. However, we limited the spaced ER experiment to only two contrasts (5 and 90%) to keep the scan time comparable to the rapid ER experiment (~7 minutes with 9 or 10 repetitions of each stimulus contrast per scan).

Our second purpose in the present study is to address the cause of nonlinearities in the BOLD timing. We did this by adapting the experiment to 7T, allowing us to use a spin echo (SE) scanning protocol, which is known to represent the signal from small vessels by decreasing the effect of large vessels (Duong et al. 2002; Yacoub et al. 2005). We were not able to rapidly scan using SE because the signal to noise ratio is very low, and the SAR would be too high. We matched the GE and SE scanning to 1.5 second TRs, using four 1.75° radius sine wave grating patches presented at 3° eccentricity.

2. Methods

3T experimental procedure and data collection

Participants

Four observers experienced in visual task fMRI experiments (3 female, 1 male) participated in this study. All participants were aware of the experimental aims; all are authors. All participants had normal or corrected-to-normal visual acuity. Participation

was voluntary and in compliance with the University of Minnesota Institutional Review Board guidelines.

Stimuli and scanning paradigm

Visual stimuli were created using Matlab and the Psychtoolbox extensions (Brainard 1997; Pelli 1997). Stimuli were presented on a NEC 2190UXi monitor with a resolution of 1024x768 pixels and a refresh rate of 60 Hz. The luminance of the monitor was calibrated using a Photo Research PR-655 photometer. The monitor had a range from 0.3 to 214 cd/m^2 with a mean gray luminance of 107 cd/m^2. The monitor was mounted to the back wall of the scanning suite; subjects viewed the monitor through a mirror mounted to the top of a clamshell 12 channel head coil.

The stimuli used for the 3T data acquisition contained contrast polarity-reversing checkerboards that reversed at a rate of 4 Hz extending across 12° of visual angle in the horizontal direction, and 9° in the vertical direction. A rapid serial visual presentation (RSVP) task was presented at the center of the screen, encouraging participants to maintain foveal fixation at the center of the screen. Letters were presented at the center of the screen at a rate of 4 Hz and subjects pressed a button every time they saw the letter 'x', which occurred with a 10% probability.

Three types of scans were run. One was a block-design checkerboard used for region of interest (ROI) localization, while the other two types were used to collect timing data: one rapid event-related (interstimulus interval of 1 second), and one spaced event-related (interstimulus interval of either 15 or 15.5 seconds). We used an m-sequence for optimal stimulus condition balancing in both types of event-related scans (Buracas and Boynton

2002). Both event-related scan types contained a 6 s 'cognitive pause' before presenting the first visual stimulus, as well as a 16 s pause after the final stimulus presentation to allow the HRF to return to baseline.

The rapid event-related experiment comprised a stimulus duration of 500 ms and an interstimulus interval of 1 second. 248 events total were displayed during each scan (48 blank, 50 each of 2.5%, 5%, 25%, and 90% contrast) for a total of 1008 TRs. Subjects performed between 4 and 8 scans per session, depending on time available and subject's ability to remain awake and attentive. Each subject performed 2 sessions.

Because of the increase in time needed for a spaced event experiment, we tested only two conditions: 5 and 90% contrast. The spaced event related experiment contained 24 events total in each scan (8 blank, 9 5% contrast, 9 90% contrast) for a total of 1528 or (4 blank, 10 5% contrast, 10 90% contrast) 1576 TRs. Subjects performed 5-6 scans per session. Each subject performed 2 sessions.

In order to independently localize the region of cortex corresponding to the retinotopic location of the checkerboard, we used a block localizer consisting of 12 seconds containing a contrast-reversing checkerboard at 90% contrast alternating against 12 seconds containing a blank, gray screen. The same RSVP task that was used for the event related experiment was used for this block localizer.

Data acquisition and pre-processing

All MR images were acquired with gradient echo EPI (TR 250ms, TE 25 ms) at 3T (Siemens TIM Trio scanner and console) using a 12-channel receive-only RF coil. Slices were oblique and perpendicular to the calcarine sulcus in the occipital cortex. The field of view was 172 x 144 with a matrix size of 64 x 48 voxels, yielding a resolution of 2.7mm

x 3mm x 2.7mm with 4 slices total. Data were reconstructed using the native Siemens software. All functional data were distortion compensated using a field map that was acquired during the functional imaging scanning session and the FUGUE toolbox distributed with FSL (http://www.fmrib.ox.ac.uk/fsl/).

Retinotopic mapping

Primary visual cortical areas (V1 and V2) were identified using standard retinotopic mapping techniques (Sereno et al. 1995; DeYoe et al. 1996; Engel et al. 1997). To map eccentricity, we used checkerboard stimuli masked by expanding and contracting annuli (checkerboards slid through the ring aperture to minimize adaptation), with rings either expanding from 0° to 11° over 24 seconds or contracting from 11° to 0° over 24 seconds. After calculating the response phase and coherence for each voxel, we used the phase values as an estimate of temporal stimulation delay and thereby identified foveal representations and approximate eccentricity representations in retinotopic areas. To identify visual area boundaries, we used a 45° sliding checkerboard wedge that rotated about the center of fixation (completing one cycle in 24 seconds) to elicit traveling waves of activity across each of the retinotopic visual areas. After similarly calculating the phase and coherence of each voxel, because the phase (or temporal delay) corresponds to the polar angle of the stimulus wedge, we estimated V1 as the area corresponding to the contralateral hemifield, and V2 as the areas corresponding to the lower contralateral quarter field dorsal to V1 and the upper contralateral quarter field ventral to V1.

fMRI data preprocessing

Alignment to volume anatomy

The first volume of the first EPI series was then aligned to a 3D MP-RAGE volume of the whole brain collected in a separate scanning session. Automatic alignment was performed after inversion of voxel intensities to match the T_1 contrast of the anatomical data (Nestares and Heeger 2000).

Segmentation of cortical surface

The gray/white matter surface was segmented and reconstructed from the MP-RAGE volume anatomy of each subject with SurfRelax (Larsson 2001). Patches of the occipital cortex were cut from the inflated whole brain and flattened to aid visualization of the functional data.

Data analysis

Data analysis was performed with code written in MATLAB. The first 8 time points (1/2 cycle) of the block experiment were discarded (none were discarded for the event-related experiment) and motion compensation was performed on the remaining data from each scan (Nestares and Heeger 2000). High pass filtering was then performed on the data with a cut-off frequency of 4 cycles per scan for the block, and 8 cycles per scan for the event related experiment.

The V1 region of interest (ROI) was defined to include voxels with stimulus coherence greater than 0.3 in the localizer scan and confined to V1 previously defined during retinotopic mapping. ROIs were determined independently in the right and left

hemispheres of each subject for each scanning session. Each experimental condition (rapid or spaced) contained estimated latencies from 16 measurements (2 hemispheres, 2 sessions, 4 subjects).

Hemodynamic response functions (HRFs) were estimated for the rapid event-related data using ordinary least squares (Hinrichs et al. 2000 Woldorff, Dale and Heinze 2000; Serences 2004). The resulting HRFs were averaged across scans, sessions, and subjects. Of note, there was no difference between the HRFs of the spaced data when they were deconvolved and then averaged (using a 12 second window after each event), or when the 15 second window after the stimulus presentation (during which there were no other stimuli presented except the RSVP task) was used (not deconvolved) and event-related averaged, suggesting that there were no amplitude nonlinearities of stimulus presentation timing in these data.

We analyzed the data by fitting the HRF with a difference of gamma functions model (least squares minimization) to the average across the entire ROI defined per subject, as well as by fitting each individual voxel within the ROI of each subject, and then averaging across the entire ROI. Fits that did not meet 'good fitting' criteria were discarded. Fits were considered 'good' if the peak time was less than 10 seconds, the full-width half-max was less than 6 seconds but greater than 0.5 seconds, and the last iteration of the fitting function changed by less than half the peak amplitude of the fitted HRF. An ROI was excluded if more than half the HRFs derived from it were poorly fitting. No ROIs were excluded. HRFs and CRFs were similar when individual voxels were fit and the average was taken in each ROI to when the voxels were averaged across an ROI and

the HRF was fit to this average. We therefore proceeded through our statistical analysis using the HRF fitted to the mean across the entire ROI (n = 16: 2 ROI per scanning session, 2 scanning sessions per subject, 4 subjects).

To quantify the timing characteristics of the HRF, we estimated the latency of onset, full-width half-max (FWHM), and time to peak. Because of uncertainty of the exact take-off point of the HRF, we used the time at which the model fit HRF first reached 10% of its peak value as an estimate of the HRF onset. To make sure that our findings were not an unique artifact of the timing values at 10% of the peak response, we also looked at the latency at 50% of the peak response. We saw the same pattern of latencies at 10% and 50% of the peak value (figure 3.3), so all subsequent analyses are reported for the values at 10% of the peak. To find the FWHM, we used the time after the peak at which the fitted HRF reached 50% of the peak amplitude minus the time before the peak at which the fitted HRF reached 50% of the peak amplitude. Time to peak was estimated as the time at which the fitted HRF reached its maximum amplitude.

7T experimental procedure and data collection

Participants

Two observers experienced in visual task fMRI experiments (both female) participated in this study. Both were aware of the experimental aims; they are both authors. All participants had normal or corrected-to-normal visual acuity. Participation was voluntary and in compliance with the University of Minnesota Institutional Review Board guidelines.

Stimuli and scanning procedure

Stimuli for this experiment were also created using Matlab and the Psychtoolbox extensions (Brainard 1997; Pelli 1997). Stimuli were projected through a waveguide from a room adjacent to the scanner room, and reflected into a mirror attached to the top of the head coil.

The stimuli used for the 7T experiment contained paired presentations of static patches of sinusoidally modulated luminance grating presented at 3° eccentricity and subtending 1.75° of visual angle. The orientation of the grating patch was random and varied with each stimulus presentation. Subjects were engaged in a 2-interval forced choice task, indicating whether one of the patches contained a contrast increment in one of the two presentation intervals.

Two types of scans were run. Block localizers (12s of stimulus, 12s of mean gray screen) were composed of rapid presentations of 2 interval forced choice trials. In one interval, all four grating patches were presented at a slightly higher contrast than during the second interval. The contrast increment varied across trials, controlled by a 3-down/l-up staircase to maintain task difficulty at a level at which each subject performed at 80% correct. There were 8 trials per block (stimulus duration = 150 ms, 150 ms blank, second stimulus duration = 150 ms, 550 ms blank, for a total trial length = 1.5 s), 10 blocks per scan, with stimuli at 90% contrast. Event related scans contained a 2-interval forced choice task in which one of the four grating patches was presented with a higher contrast (increment varied across trials) in only one of the intervals. The pedestal contrast for all the patches in both intervals varied between four contrast levels (5, 10, 30, 90%). Each

trial contained a first stimulus presentation (250 ms), a blank screen (200 ms), a second stimulus presentation (250 ms), and a blank screen for a variable length of time with an average inter-trial interval of 4.5 s (randomly selected from 3, 4.5 or 6 s with a uniform distribution). Trial contrast was randomized. Separate 3-down/1-up staircases controlled task difficulty for each contrast level. The first stimulus occurred 6 s after the start of the scan, and a mean gray screen was presented for 16 s after the final stimulus presentation to allow the HRFs to return to baseline. Scans comprised 15 presentations of each contrast level, for a total of 188 to 212 TRs.

Data acquisition and pre-processing

All MR images were acquired with gradient echo EPI (TR 1.5s, TE 20 ms) or with spin echo EPI (TR 1.5s, TE 40ms) at 7T (Siemens TIM Trio scanner and console) using a surface coil. Slices were oblique and perpendicular to the calcarine sulcus in the occipital cortex. The field of view was 192 x 96 with a matrix size of 96 x 48 voxels, yielding a resolution of 2mm x 2mm x 2mm with 12 slices total. Data were reconstructed using the native Siemens software. All functional data were distortion compensated using a field map that was acquired during the functional imaging scanning session and the FUGUE toolbox distributed with FSL (http://www.fmrib.ox.ac.uk/fsl/).

Retinotopic mapping and fMRI data preprocessing

Retinotopic mapping (performed in a separate scanning session), as well as alignment of the volume anatomy and segmentation was identical to that described for the 3T data.

The motion compensation process was also identical to that used for the 3T data. The data were high pass filtered with an 8 cycle per scan cut off.

Data analysis

Data analysis was performed with the same code written in MATLAB that was used for 3T data analysis. Regions of interest were defined by voxels that were both within the retinotopically defined V1, and showed a high coherence (0.3) to target event related scans.

The first 8 time points (1/2 cycle) of the target localizer block experiment were discarded (none were discarded for the event-related experiment) and motion compensation was performed on the remaining data from each scan (Nestares and Heeger 2000). High pass filtering was then performed on the data with a cut-off frequency of 4 cycles per scan for the block, and 8 cycles per scan for the event related experiment.

The V1 region of interest (ROI) was defined to include voxels with stimulus coherence greater than 0.3 in the localizer scan and confined to V1 previously defined during retinotopic mapping. ROIs were determined independently in the regions corresponding to the four sine-wave grating patches that were presented, two patches in each the right and left hemispheres of each subject (see figure 3.1, c and d). Each data type (GE or SE) contained estimated latencies from 8 measurements (4 grating patch ROIs, 1 session, 2 subjects).

HRFs were estimated for the rapid event-related data using ordinary least squares (Hinrichs et al. 2000 Woldorff, Dale and Heinze 2000; Serences 2004). The resulting HRFs were averaged across scans and subjects. Both the GE and SE event-related data

were analyzed using deconvolution using ordinary least squares and event-related averaging.

We analyzed the data by fitting the difference-of-gamma functions model to the average HRF across the entire ROI defined per subject, as well as by fitting each individual voxel within the ROI of each subject, and then averaging fit parameters across the entire ROI. Because the fit to the average through subjects was similar to the average of fits to each subject, we used the fit to the average through subjects for our analyses. We excluded poorly fit HRFs using the same criteria as the 3T protocol. All ROI-average fits from Subject 1 were thereby excluded. The data from this subject exhibited dramatic post-stimulus undershoots which our fitting function did not handle appropriately. We therefore used the HRF average across estimated voxel fits, rejected poorly fitting voxels, and estimated timing parameters from voxel fit HRFs. This revealed a much higher apparent fitting success (on visual inspection of the fits to the data) and smaller variance in the timing characteristics and peak amplitude estimates, suggesting that these fits were more accurate. Because we have hundreds of voxels and taking the average and standard error of the mean over all voxels would give us an unreasonably small standard error of the mean because it would incorrectly assume voxel independence, we first averaged over the voxels within an ROI and then took the average and standard error of the mean across all ROIs. We proceeded through our statistical analysis using the mean of the HRFs fitted to individual voxels (n = 8: 4 ROIs per scanning session, 1 scanning session per subject, 2 subjects).

To quantify the timing characteristics of the HRF, we estimated the latency of onset, full-width half-max (FWHM), and time to peak with identical parameters as used in the 3T experiment. Statistical tests of latency were done only with the latencies estimated from 10% peak amplitude, as was also done at 3T because the 7T data also show a similar trend in latency estimates between 10% and 50% of the peak, as shown in figure 3.5a and 3.5c.

We attempted to determine whether latency was dependent on the size of the vessel. Given that signals from large draining veins generally have a higher variance though time than do signals from small vessels (de Zwart et al. 2005 Kellman, Fukunaga, Chu, Koretsky, Frank, and Duyn 2005; Olman et al. 2007), if large veins were the cause of timing differences, we expected higher variance voxels to show a bigger effect of contrast on latency. We compared the latencies of the voxel responses when the data were broken into groups by voxel variance bins. Variance was calculated for each voxel through time, and all voxels were sorted. Voxels were assigned into one of three bins (lowest, middle, or top third by variance) (figure 3.6). We did not, however, see any latency dependence on variance.

3. Results

Data collected

Two sessions of the 3T data were run on each subject for each condition (rapid and spaced event related), for a total of 4 scanning sessions per subject. Between 4 and 6 scans were completed during each session. One session of the 7T data was collected, 2 subjects, 4 GE and 4 SE scans per subject.

3T Data

We observed the expected HRF amplitude modulation with increasing contrast in both the rapid and spaced ER experiments. Individual subject HRFs are shown in figure 3.2. Peak amplitude as a function of contrast (the contrast response function) showed saturation at high contrast, which is consistent with contrast response functions estimated from psychometric data (Boynton 1999) (figure 3.3c,f).

The mean timing characteristic estimates are shown in table 3.1 for the rapid ER experiment, and table 3.2 for the spaced ER experiment. To compare differences in each timing characteristic as a function of contrast, we performed paired t-tests. To estimate a difference in trend, we compared the average over the high contrast to the average over the lower contrast responses. For the rapid ER, this was a comparison of the average response to 2.5 and 5% contrast stimuli against the average response to 25 and 90% contrast stimuli. For the spaced ER, there was no averaging because there were only two contrasts. Asterisks denote a significant difference between groups, and are only presented in tables 3.1 and 3.2.

We found a significant effect of contrast on latency. The trend between latency estimated by 10% of the peak and 50% of the peak were similar (figure 3.3a and 3.3c). All the statistical tests reported were done on the latency estimated at the time the peak reached 10% of the peak amplitude. For the rapid ER latency data, there were no significant differences between the two low contrast responses or between the two high contrast responses (2.5 and 5%, or 25 and 90%), but both low contrast levels were different from both high contrast levels. A corollary finding was a significant difference

between the mean over the low contrast compared to the mean over the high contrast latencies (depicted by * in table 3.1). The spaced ER latency data showed the same trend, with a significant difference between the low (5%) and high (90%) contrast HRF latency.

There was also a weak effect of contrast on HRF width. For the rapid ER experiment, there were no significant differences between any of the individual FWHM estimates, nor between the mean over the low contrast FWHM compared to the mean over the high contrast. However, the spaced ER experiment did reveal a significant difference between the low contrast FWHM compared to the high contrast FWHM ($p = 0.0382$).

We did not find an effect of contrast on time to peak. For both the rapid and spaced ER experiments, there were no significant group differences between any contrast response peak time pair (2.5 to 5%, 2.5% to 25%, etc) nor between the mean over the high contrast time to peak compared to the low contrast time to peak.

7T Data

Four ROIs (one in each visual quadrant) were analyzed for each subject. We found similar results to the 3T data in both the GE and SE 7T data. The data show the expected HRF saturating nonlinear increase in peak amplitude as a function of stimulus contrast (figure 3.4). As with the 3T data, we estimated the mean latency, FWHM, and time to peak for each ROI and performed paired t-tests for each combination of stimulus contrast HRFs, as well as paired t-tests comparing the average low versus high contrast timing parameter estimates, for both the GE and SE data. Timing characteristics are provided in tables 3.3 and 3.4 as mean across ROI ± standard deviation. These timing characteristics area also displayed in figure 3.5 (errorbars are standard error of the mean, $n = 8$).

Latency varied significantly with contrast for the GE but not the SE data. Although this finding was weaker in the 7T data than the 3T data, the mean low contrast HRF estimated latency was significantly different from the high contrast HRF estimated latency. However, this appeared to be driven primarily by the difference between the 5% and 90% contrast stimuli, as these were the only contrasts that showed a difference before averaging ($p < 0.05$). However, the 5% and 30% contrast latencies almost reached significance ($p = 0.0742$) as well. The SE data, conversely, showed no significant differences between latencies at any contrast or over low versus high contrast averages, and nothing even nearing significance (all $p > 0.1$).

Unlike the 3T data, the FWHM did not appear to vary with contrast, and no significance testing rejected the null hypothesis of equal widths. FWHM at 7T both using GE and SE techniques did not appear to change with contrast (although they are different from each other).

Also unlike the 3T data, there were some indications that the peak time at 7T varied with contrast. Time to peak for 5% contrast was significantly different from time to peak for 10% contrast in the GE data, and time to peak was different for the following comparisons of the SE data: 5 to 30%, 5 to 90%, 10 to 30%, 10 to 90%, and the average low to the average high contrast. Similar to the observation at 3T, variance binning (like that performed on the 3T data) did not reveal a latency correlation with variance in either the GE or SE data acquired at 7T (figure 3.6).

4. Discussion

Difference in latency as a function of contrast

Our results suggest that timing nonlinearities are present in event-related experiments at 3T regardless of whether stimuli are rapidly presented or are widely spaced. This is not reported in previous studies of either timing profiles or nonlinearities in the BOLD HRF. The fact that this has not been observed before may be due to the lower temporal resolution of previous measurements of the contrast-dependent HRF, or the fact that studies generally analyze the peak stimulus response time, which we found to be invariant in time with respect to contrast (or peak amplitude) in either a spaced or rapid event related stimulus design at 3T. This finding was consistent across GE data obtained at high field strength (7T), but not found in SE data at 7T. This raises the possibility that latency varies with vessel size. However, when attempting to sort latency by vessel size (by binning into three voxel variance levels, which roughly correlates with vessel size) we found no difference in the three variance bins.

Peak broadening

Because we found a latency dependence on contrast, we measured the FWHM of the HRF as a representative of HRF shape. In the 3T data, we found a significant effect of contrast on the FWHM. In other words, the peak appeared to broaden with increased contrast. Although this does not explain why we would see an early onset with higher contrast, it does suggest that contrast changes the shape of the HRF in a more complex way than just shifting the onset time, or widening the HRF. This provides further

evidence that a better fit to the data may provide more insight into the timing of the hemodynamic response, and that a simple gain or shift will not correct timing parameter estimates.

Time to peak shifts

Counterintuitively, we measured a significant effect of contrast on time to peak at 7T for both GE and SE data, but no such effect on rapid or spaced presentation data at 3T. One possible cause for this may be due to the combination of low contrast to noise ratio (CNR - for the SE data only) at 7T, and individual voxel fitting and exclusion that was performed (as previously described) at 7T but not at 3T. Because voxels with low CNR were more likely to be excluded due to poor fitting, this would leave only strongly modulated and/or less noisy voxels for analysis. Because of this, we may expect either to see modulation mostly due to larger vessels, which produce stronger modulation, or more accurate timing parameter estimates. However, given the very fast time to peak estimates (~1.5 sec) we measured for the SE data, this is highly unlikely to be physiologically plausible, or the correct reason for our observation. It is more likely the following. We used the same ROI for both the GE and SE data analysis so that we could directly compare the voxel responses between the two (since the data were collected in each subject during the same scanning session). However, the GE ROI was used, which included voxels that would not be significantly modulated during the SE scans. This is especially true at low contrast, where the signal is small, and, for SE data, the CNR is very low. We estimated the timing characteristics from fits to the sampled data, the lower CNR did result in poorer fits, and more discarded fits. These poorer fits seemed to

estimate the HRF peak to be later in the time course (more frequent in lower contrast stimuli), and may be the reason for the observed time to peak dependence on contrast. Therefore, the time to peak contrast dependency we observed is likely an artifact of GE ROI definition in SE acquired data, or low CNR, either of which could result in poor fitting. Although CNR may have affected fitting in the SE data, there is no reason to believe this problem occurred in the much stronger modulated GE data (at either 3T or 7T).

Determining onset time

Because of the 250 ms TR of our 3T data, although rapidly sampled, reduces SNR by approximately 40%, creating noise and making a baseline estimate and a true estimate of the HRF onset from the noise remained difficult. Because of this, we fit the HRFs with gamma functions. On the other hand, our 7T data appeared less noisy but our sampling rate was slow, so we also fit gamma functions to estimate the HRF for these data. Using gamma functions constrained us by the assumption that this function is in fact an appropriate estimate of the HRF. Because our purpose was to determine whether nonlinearities in the timing were present, and these nonlinearities may be the result of deviations from a function well fit by the gamma function, we feel that another strategy may reveal better estimates, and perhaps a clearer story of the relationship between latency and stimulus contrast. We did attempt to blur the real data in order to smooth the HRF, but previous studies have already shown that blurring through time is analogous to undersampling, benefiting the determination of timing latency no better than a slow TR (Dilharreguy et al. 2003). We therefore proceeded only with HRFs fitted by gamma

functions with the understanding that timing estimates may be better determined by a less constrained model.

Mechanisms of timing nonlinearities

One possible mechanism may be a pooled response over both slow but nonspecific response from large vessels with the faster, stimulus specific capillary response (Hulvershorn et al. 2005). If the more salient or higher contrast stimuli evoke a response requiring more large vein response, then high contrast stimuli may contain a different proportion of small to large vessel contributions, which may in turn contribute responses with different temporal dynamics. At low contrast, more capillary response could contribute hemodynamics with a slower response, while at high contrast, more large vein response could contribute hemodynamics with a faster response.

5. Conclusion

Finding timing latency differences between low and high contrast stimuli has implications for the use of fMRI BOLD data in drawing relationships between data acquired at low resolution and moderate field strength for use in making determinations about the neural response timing in cortex, as is done in dynamic causal modeling. The BOLD latency nonlinearities that we have found are not likely to be due to neural events since the observed BOLD latency nonlinearities are on the order of 100's of milliseconds, unlike neural contrast nonlinearities, which occur in the 10's of milliseconds. This suggests a vascular mechanism. Because we also observed contrast dependent latency at 7T in GE acquired data, but not in SE acquired data, our findings support this hypothesis.

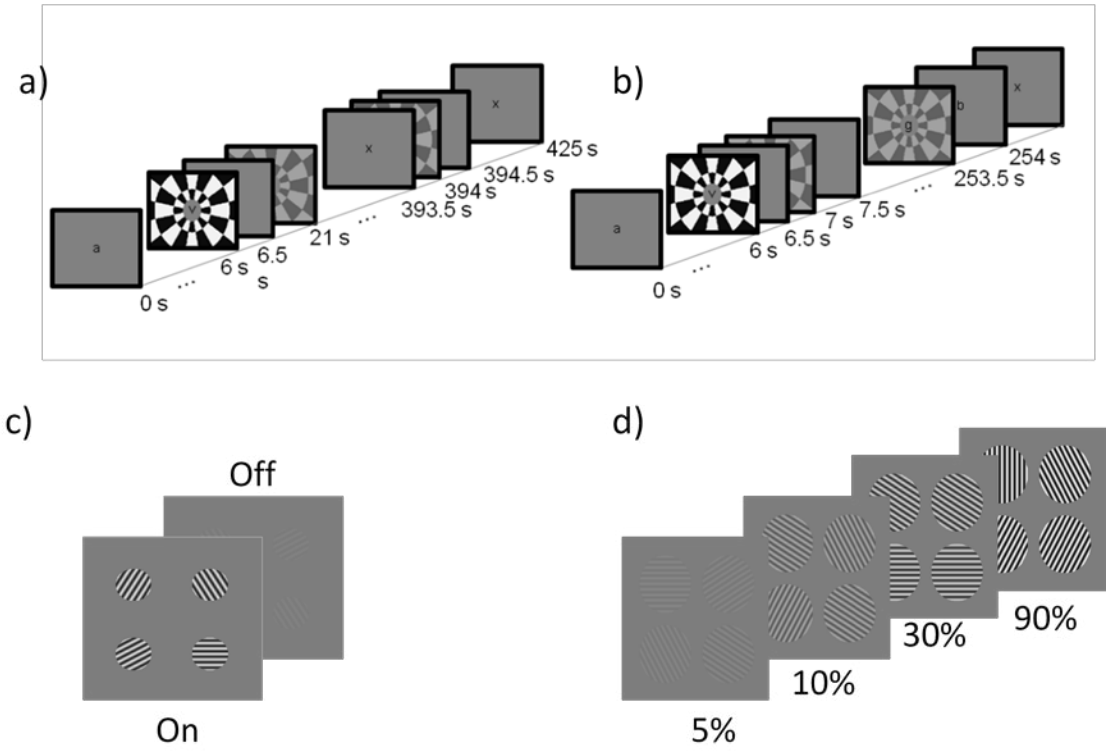

Figure 3.1. Stimulus paradigm for latency experiment. (a and b) show 3T rapid (a) and spaced (b) experiments. (c and d) show paradigms for the 7T experiment: c) target localizer 'on' and 'off' blocks; d) the 4 contrast levels (left to right = 5, 10, 30, 90%) presented in an ER experiment scanning with both GE and SE paradigm.

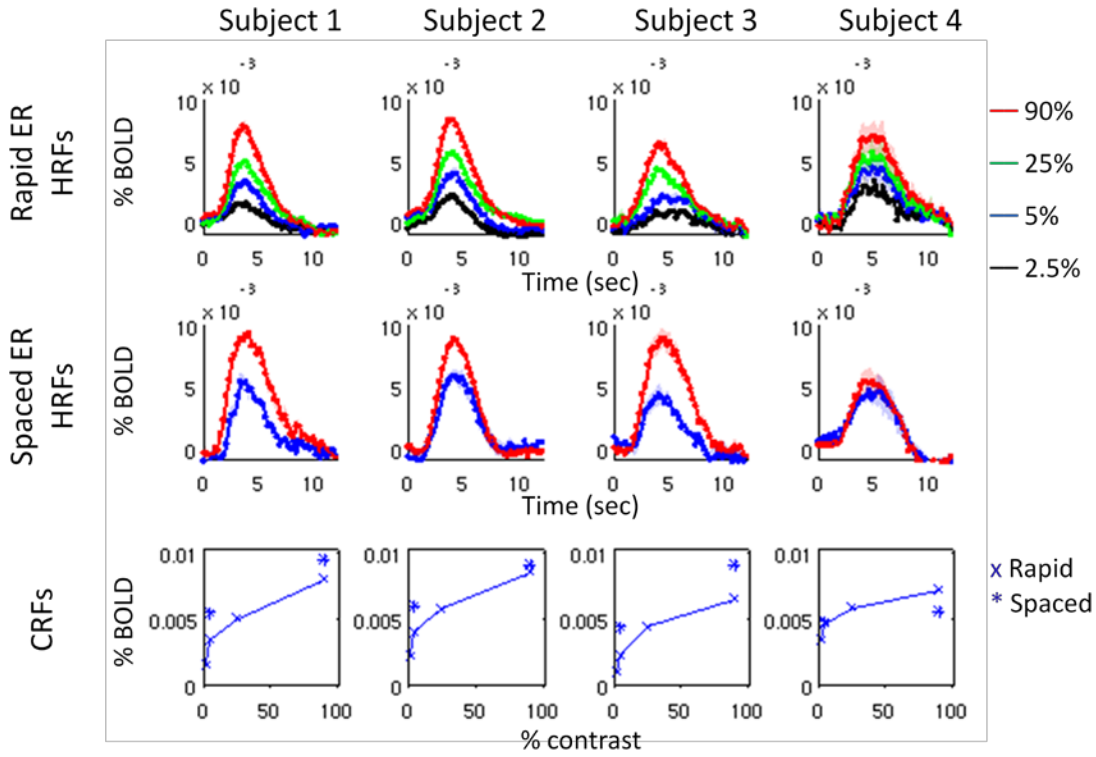

Figure 3.2. 3T individual subject average HRFs and CRFs. Row 1) Rapid event related, Row 2) Spaced event related.. Row 3) CRFs. 'x' denotes rapid ER, '*' denotes spaced ER peak amplitude vs. % contrast.

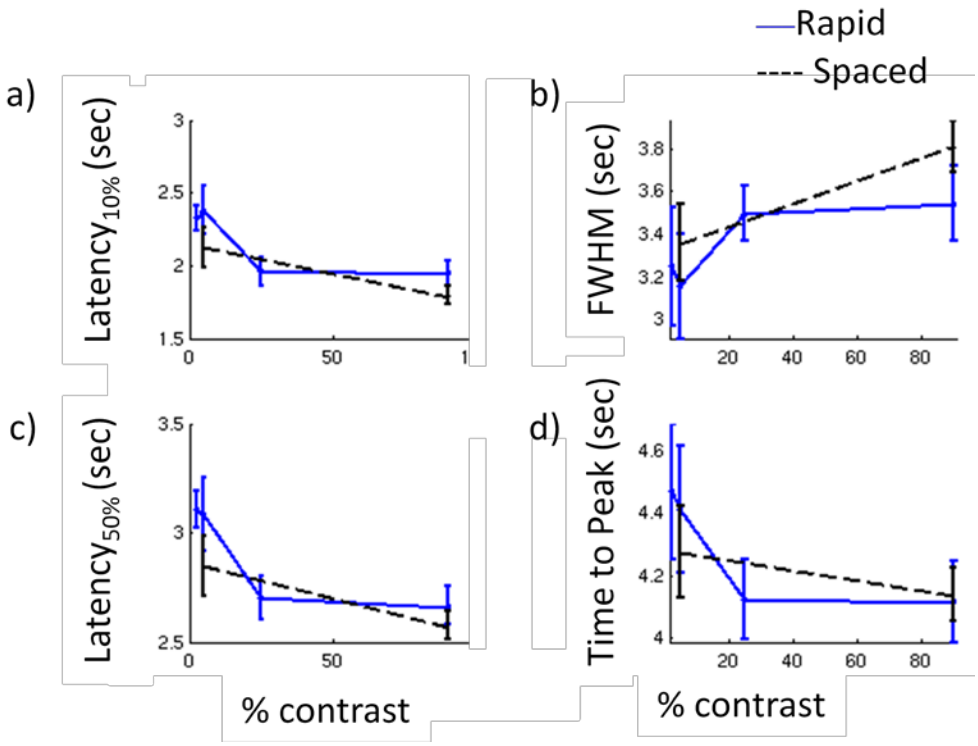

Figure 3.3. Latency, FWHM, and Peak times (3T). Timing parameters were estimated from difference of gamma functions model fit to HRFs for 3T experiment. Latency at 10% peak (a), latency at 50% peak (c), FWHM (b), and Time to Peak (d) vs. % contrast for the 3T experiment. Rapid ER (solid blue line) and Spaced ER (dashed black line) experiments. Data were averaged over all ROIs in all subjects (16 total = 2 ROIs per scanning session, 2 scanning sessions per subject, 4 subjects). Error bars denote SEM (n=16).

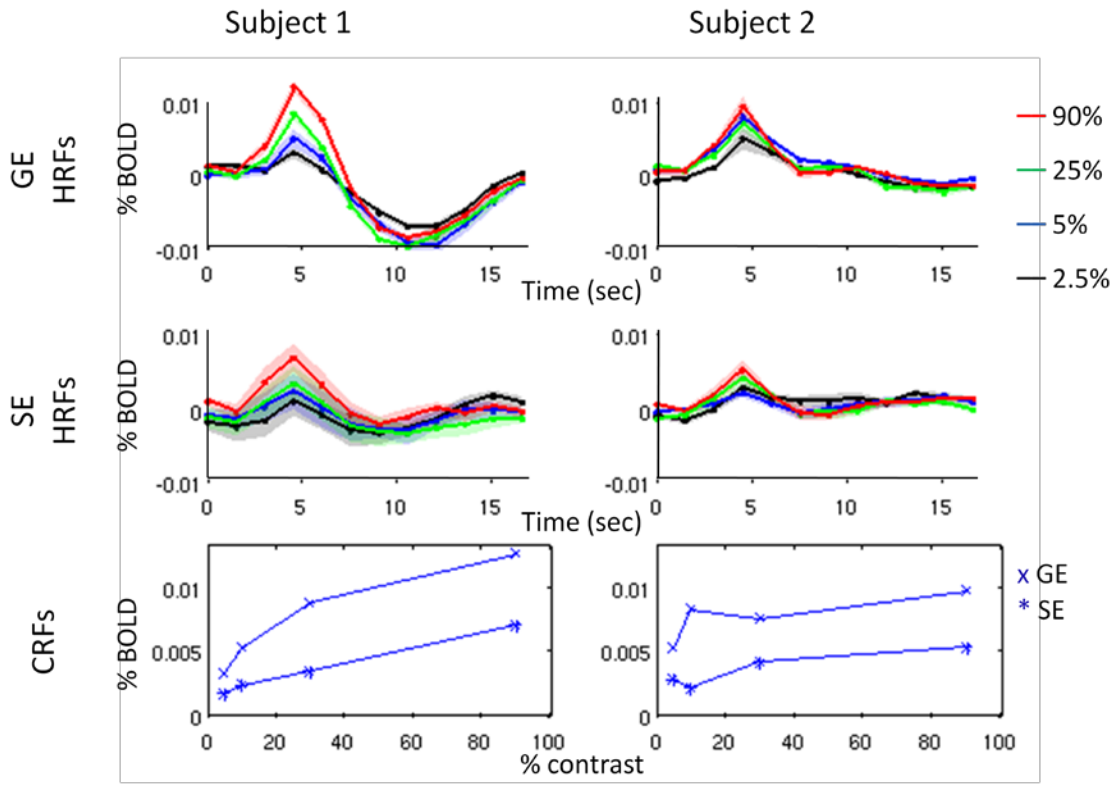

Figure 3.4. 7T Individual subject average HRFs and CRFs. Row 1) gradient echo average HRF for each subject. Row 2) spin echo average HRF for each subject. Row 3) CRF. 'x' denotes GE, '*' denotes SE peak amplitudes vs. % contrast.

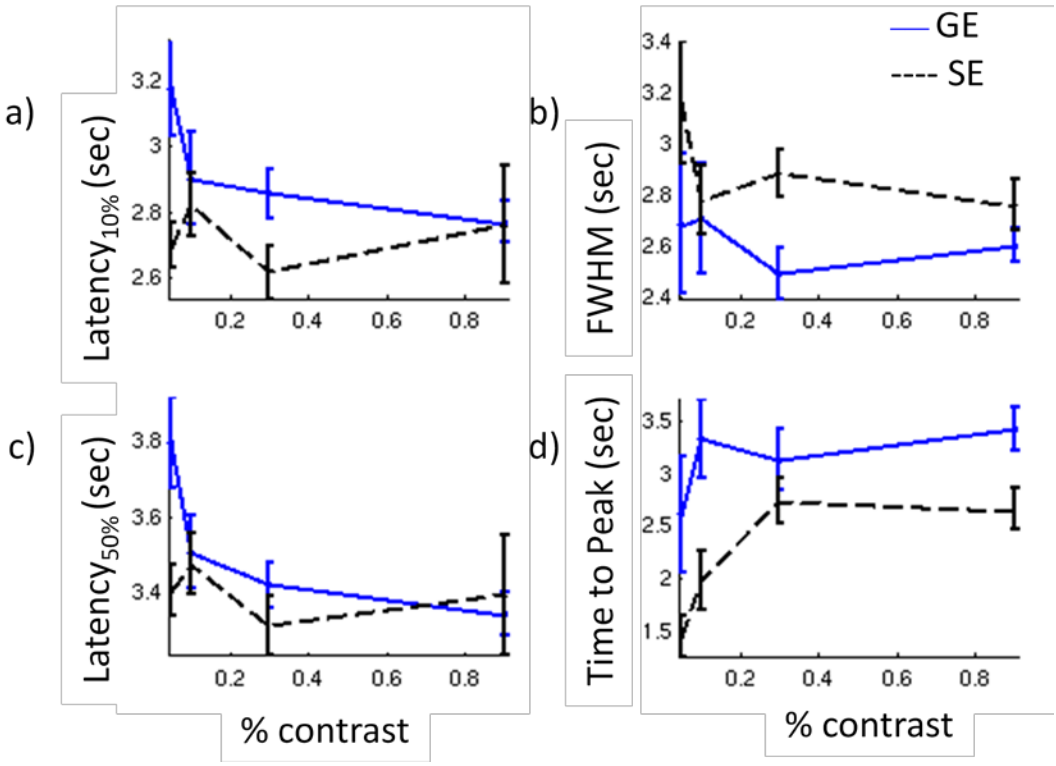

Figure 3.5. Latency, FWHM, and Peak times (7T). Timing parameters were estimated from difference of gamma functions model fit to HRFs for 7T experiment. Latency at 10% peak (a), latency at 50% peak (c), FWHM (b), and Time to Peak (d) vs. % contrast for the 7T experiment. GE (solid blue line) and SE (dashed black line) experiments. Data were averaged over all ROIs in all subjects; timing estimated from HRF fit to each ROI (8 total = 4 ROIs per scanning session, 1 scanning sessions per subject, 2 subjects). Error bars denote SEM (n=8).

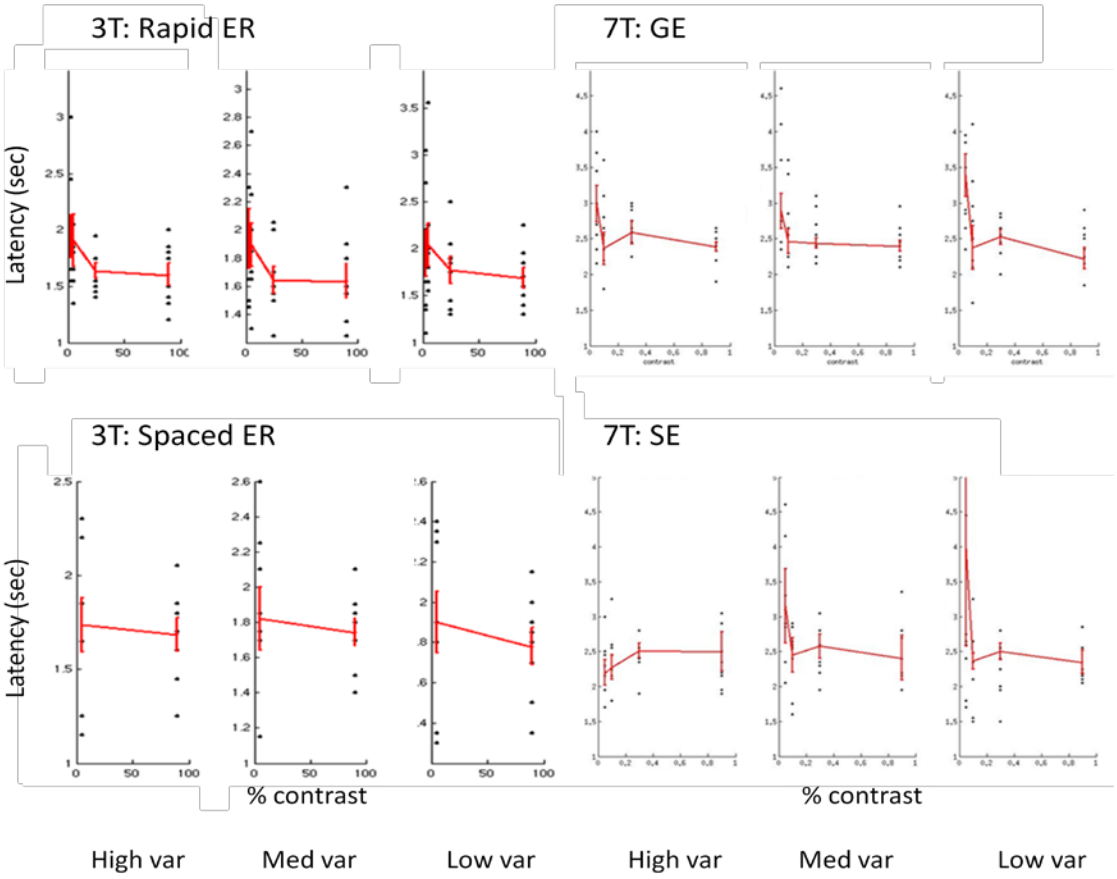

Figure 3.6. Latency binned by voxel variance. Left 3 columns show 3T data; Right 3 columns show 7T data.

Table 3.1: 3T Rapid ER timing

	Latency	FWHM	Time to Peak
2.5%	2.33 ± 0.335	3.25 ± 1.11	4.47 ± 0.877
5%	2.38 ± 0.668	3.16 ± 0.973	4.41 ± 0.815
25%	1.96 ± 0.392	3.50 ± 0.503	4.12 ± 0.508
90%	1.95 ± 0.356	3.54 ± 0.702	4.12 ± 0.525
Lo contrast avg	2.35 ± 0.443	3.20 ± 0.955	4.44 ± 0.790
Hi contrast avg	1.95 ± 0.285	3.52 ± 0.545	4.12 ± 0.477
Lo/Hi difference	0.397 ± 0.0486*	-0.316 ± 0.793	0.323 ± 0.658

Reported as: mean ± sd

Table 3.2: 3T Spaced ER timing

	Latency	FWHM	Time to Peak
5%	2.13 ± 0.544	3.36 ± 0.714	4.28 ± 0.601
90%	1.80 ± 0.257	3.81 ± 0.497	4.14 ± 0.351
Lo/Hi difference	0.331 ± 0.554*	-0.447 ± 0.787*	0.137 ± 0.478

Reported as: mean ± sd

Table 3.3: 7T GE ER timing

	Latency	FWHM	Time to Peak
5%	3.18 ± 0.445	2.69 ± 0.821	2.61 ± 1.66
10%	2.90 ± 0.422	2.71 ± 0.652	3.34 ± 1.13
30%	2.86 ± 0.230	2.50 ± 0.295	3.13 ± 0.879
90%	2.77 ± 0.186	2.61 ± 0.200	3.42 ± 0.627
Lo contrast avg	3.04 ± 0.382	2.70 ± 0.707	2.97 ± 1.37
Hi contrast avg	2.81 ± 0.186	2.55 ± 0.182	3.27 ± 0.694
Lo/Hi difference	0.227 ± 0.262*	0.149 ± 0.552	-0.302 ± 1.05

Reported as: mean ± sd

Table 3.4: 7T SE ER

	Latency	FWHM	Time to Peak
5%	2.70 ± 0.210	3.16 ± 0.729	1.45 ± 0.572
10%	2.82 ± 0.287	2.78 ± 0.406	1.98 ± 0.844
30%	2.62 ± 0.242	2.89 ± 0.279	2.74 ± 0.629
90%	2.76 ± 0.545	2.76 ± 0.295	2.66 ± 0.605
Lo contrast avg	2.76 ± 0.219	2.97 ± 0.473	1.72 ± 0.601
Hi contrast avg	2.69 ± 0.319	2.83 ± 0.204	2.70 ± 0.585
Lo/Hi difference	0.0703 ± 0.340	0.146 ± 0.421	-0.983 ± 0.283*

Reported as: mean ± sd

Literature Cited

Albrecht, D. G. (1995). "Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions." Vis Neurosci **12**(6): 1191-210.

Albrecht, D. G., W. S. Geisler, et al. (2002). "Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function." J Neurophysiol **88**(2): 888-913.

Appelbaum, L. G., A. R. Wade, et al. (2008). "Figure-ground interaction in the human visual cortex." J Vis **8**(9): 8 1-19.

Beck, J. (1966). "Effect of surround size on the perception of texture patterns." J Exp Psychol **72**(1): 68-75.

Beck, J. (1983). "Textural segmentation, second-order statistics, and textural elements." Biol Cybern **48**(2): 125-30.

Beck, J., A Sutter, and R Ivry (1987). "Spatial frequency channels and perceptual grouping in texture segregation." Computer Vision, Graphics, and Image Processing **37**(2): 299-325.

Bergen, J. R. and E. H. Adelson (1988). "Early vision and texture perception." Nature **333**(6171): 363-4.

Birn, R. M., Z. S. Saad, et al. (2001). "Spatial heterogeneity of the nonlinear dynamics in the FMRI BOLD response." Neuroimage **14**(4): 817-26.

Boyaci, H., F. Fang, et al. (2007). "Responses to lightness variations in early human visual cortex." Curr Biol **17**(11): 989-93.

Boynton, G. M., S. A. Engel, et al. (1996). "Linear systems analysis of functional magnetic resonance imaging in human V1." J Neurosci **16**(13): 4207-21.

Brainard, D. H. (1997). "The Psychophysics Toolbox." Spat Vis **10**(4): 433-6.

Brown, B. D. a. C. (1982). Ballard and Brown's Computer Vision. New York, Prentice Hall.

Buckner, R. L. (1998). "Event-related fMRI and the hemodynamic response." Hum Brain Mapp **6**(5-6): 373-7.

Buckner, R. L., P. A. Bandettini, et al. (1996). "Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging." Proc Natl Acad Sci U S A **93**(25): 14878-83.

Budd, T. W., R. J. Barry, et al. (1998). "Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness." Int J Psychophysiol **31**(1): 51-68.

Buracas, G. T. and G. M. Boynton (2002). "Efficient design of event-related fMRI experiments using M-sequences." Neuroimage **16**(3 Pt 1): 801-13.

Burgess, A. E., X. Li, et al. (1997). "Visual signal detectability with two noise components: anomalous masking effects." J Opt Soc Am A Opt Image Sci Vis **14**(9): 2420-42.

Burock, M. A., R. L. Buckner, et al. (1998). "Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI." Neuroreport **9**(16): 3735-9.

Burr, D. C., M. C. Morrone, et al. (1992). "Electro-physiological investigation of edge-selective mechanisms of human vision." Vision Res **32**(2): 239-47.

Buxton, R. B., E. C. Wong, et al. (1998). "Dynamics of blood flow and oxygenation changes during brain activation: the balloon model." Magn Reson Med **39**(6): 855-64.
Caelli, T., B. Julesz, et al. (1978). "On perceptual analyzers underlying visual texture discrimination: Part II." Biol Cybern **29**(4): 201-14.
Chubb, C., M. S. Landy, et al. (2004). "A visual mechanism tuned to black." Vision Res **44**(27): 3223-32.
Cornelissen, F. W., A. R. Wade, et al. (2006). "No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex." J Neurosci **26**(14): 3634-41.
Cowey, A. and E. T. Rolls (1974). "Human cortical magnification factor and its relation to visual acuity." Exp Brain Res **21**(5): 447-454.
Dale, A. M., R. L. Buckner (1997). "Selective Averaging of Rapidly Presented Individual Trials Using fMRI." Human Brain Mapping **5**: 329-340.
David, S. V., W. E. Vinje, et al. (2004). "Natural stimulus statistics alter the receptive field structure of v1 neurons." J Neurosci **24**(31): 6991-7006.
de Zwart, J. A., A. C. Silva, et al. (2005). "Temporal dynamics of the BOLD fMRI impulse response." Neuroimage **24**(3): 667-77.
de Zwart, J. A., P. van Gelderen, et al. (2009). "Hemodynamic nonlinearities affect BOLD fMRI response timing and amplitude." Neuroimage **47**(4): 1649-58.
Dean, A. F. and D. J. Tolhurst (1986). "Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat striate cortex." Exp Brain Res **62**(1): 143-51.
DeYoe, E. A., G. J. Carman, et al. (1996). "Mapping striate and extrastriate visual areas in human cerebral cortex." Proc Natl Acad Sci U S A **93**(6): 2382-6.
Dilharreguy, B., R. A. Jones, et al. (2003). "Influence of fMRI data sampling on the temporal characterization of the hemodynamic response." Neuroimage **19**(4): 1820-8.
Duong, T. Q., E. Yacoub, et al. (2002). "High-resolution, spin-echo BOLD, and CBF fMRI at 4 and 7 T." Magn Reson Med **48**(4): 589-93.
Duvernoy, H. M., S. Delon, et al. (1981). "Cortical blood vessels of the human brain." Brain Res Bull **7**(5): 519-79.
Engel, S. A., G. H. Glover, et al. (1997). "Retinotopic organization in human visual cortex and the spatial precision of functional MRI." Cereb Cortex **7**(2): 181-92.
Fang, F., H. Boyaci, et al. (2009). "Border ownership selectivity in human early visual cortex and its modulation by attention." J Neurosci **29**(2): 460-5.
Fogel I, a. D. S. (1989). "Gabor filters as texture discriminator." Biological Cybernetics **61**(2): 103-113.
Friston, K. J., D. E. Glaser, et al. (2002). "Classical and Bayesian inference in neuroimaging: applications." Neuroimage **16**(2): 484-512.
Friston, K. J., L. Harrison, et al. (2003). "Dynamic causal modelling." Neuroimage **19**(4): 1273-302.
Gawne, T. J., T. W. Kjaer, et al. (1996). "Latency: another potential code for feature binding in striate cortex." J Neurophysiol **76**(2): 1356-60.

Geisler, W. S. and K. L. Chou (1995). "Separation of low-level and high-level factors in complex tasks: visual search." Psychol Rev **102**(2): 356-78.
Glover, G. H. (1999). "Deconvolution of impulse response in event-related BOLD fMRI." Neuroimage **9**(4): 416-29.
Grinband, J., T. D. Wager, et al. (2008). "Detection of time-varying signals in event-related fMRI designs." Neuroimage **43**(3): 509-20.
Haddon, J., and JF Boyce (1990). "Image segmentation by unifying region and boundary information." IEEE Transactions of Pattern Analysis and Machine Intelligence **12**(10): 929-948.
Halpern, D. L. and R. R. Blake (1988). "How contrast affects stereoacuity." Perception **17**(4): 483-95.
Harrison, S. J. and D. R. Keeble (2008). "Within-texture collinearity improves human texture segmentation." Vision Res **48**(19): 1955-64.
Haynes, J. D., R. B. Lotto, et al. (2004). "Responses of human visual cortex to uniform surfaces." Proc Natl Acad Sci U S A **101**(12): 4286-91.
Heckman, G. M., S. E. Bouvier, et al. (2007). "Nonlinearities in rapid event-related fMRI explained by stimulus scaling." Neuroimage **34**(2): 651-60.
Hegde, J. and D. C. Van Essen (2007). "A comparative study of shape representation in macaque visual areas v2 and v4." Cereb Cortex **17**(5): 1100-16.
Hinrichs, H., M. Scholz, et al. (2000). "Deconvolution of event-related fMRI responses in fast-rate experimental designs: tracking amplitude variations." J Cogn Neurosci **12 Suppl 2**: 76-89.
Hulvershorn, J., L. Bloy, et al. (2005). "Temporal resolving power of spin echo and gradient echo fMRI at 3T with apparent diffusion coefficient compartmentalization." Hum Brain Mapp **25**(2): 247-58.
Hurlbert, A. C. and T. A. Poggio (1988). "Synthesizing a color algorithm from examples." Science **239**(4839): 482-5.
Julesz, B. (1981). "Textons, the elements of texture perception, and their interactions." Nature **290**(5802): 91-7.
Julesz, B. (1986). "Texton gradients: the texton theory revisited." Biol Cybern **54**(4-5): 245-51.
Julesz, B. and T. V. Papathomas (1984). "On spatial-frequency channels and attention." Percept Psychophys **36**(4): 398-9.
Kourtzi, Z. and E. Huberle (2005). "Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation." Neuroimage **28**(2): 440-52.
Lamme, V. A. (1995). "The neurophysiology of figure-ground segregation in primary visual cortex." J Neurosci **15**(2): 1605-15.
Lamme, V. A., B. W. van Dijk, et al. (1993). "Organization of texture segregation processing in primate visual cortex." Vis Neurosci **10**(5): 781-90.
Lamme, V. A., B. W. van Dijk, et al. (1993). "Contour from motion processing occurs in primary visual cortex." Nature **363**(6429): 541-3.
Landy, M. S. and J. R. Bergen (1991). "Texture segregation and orientation gradient." Vision Res **31**(4): 679-91.

Larsson (2001). "Imaging Vision: Functional Mapping of Intermediate Visual Processes in Man."
Larsson, J., M. S. Landy, et al. (2006). "Orientation-selective adaptation to first- and second-order patterns in human visual cortex." J Neurophysiol **95**(2): 862-81.
Legge, G. E. and J. M. Foley (1980). "Contrast masking in human vision." J Opt Soc Am **70**(12): 1458-71.
Levitt, H. (1971). "Transformed up-down methods in psychoacoustics." J Acoust Soc Am **49**(2): Suppl 2:467+.
Li, W., V. Piech, et al. (2006). "Contour saliency in primary visual cortex." Neuron **50**(6): 951-62.
Li, W., V. Piech, et al. (2008). "Learning to link visual contours." Neuron **57**(3): 442-51.
Lindgren, J. T., J. Hurri, et al. (2008). "Spatial dependencies between local luminance and contrast in natural images." J Vis **8**(12): 6 1-13.
Lindquist, M. A. and T. D. Wager (2007). "Validity and power in hemodynamic response modeling: a comparison study and a new approach." Hum Brain Mapp **28**(8): 764-84.
Lindquist, M. A., C. Waugh, et al. (2007). "Modeling state-related fMRI activity using change-point theory." Neuroimage **35**(3): 1125-41.
Liu, H. and J. Gao (2000). "An investigation of the impulse functions for the nonlinear BOLD response in functional MRI." Magn Reson Imaging **18**(8): 931-8.
Luntinen, O., J. Rovamo, et al. (1995). "Modelling the increase of contrast sensitivity with grating area and exposure time." Vision Res **35**(16): 2339-46.
Mandeville, J. B., J. J. Marota, et al. (1998). "Dynamic functional imaging of relative cerebral blood volume during rat forepaw stimulation." Magn Reson Med **39**(4): 615-24.
Marreiros, A. C., S. J. Kiebel, et al. (2008). "Dynamic causal modelling for fMRI: a two-state model." Neuroimage **39**(1): 269-78.
Martin, D., C Fowlkes, D Tal, J Malik (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Proceedings of the 8th IEEE International conference on computer vision. Los Alamitos, CA, IEEE Computer Society Press**:** 416-425.
McAnany, J. J. and K. R. Alexander (2009). "Contrast thresholds in additive luminance noise: Effect of noise temporal characteristics." Vision Res **49**(11): 1389-96.
Menon, R. S., D. C. Luknowsky, et al. (1998). "Mental chronometry using latency-resolved functional MRI." Proc Natl Acad Sci U S A **95**(18): 10902-7.
Miezin, F. M., L. Maccotta, et al. (2000). "Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing." Neuroimage **11**(6 Pt 1): 735-59.
Moller, P. and A. C. Hurlbert (1996). "Psychophysical evidence for fast region-based segmentation processes in motion and color." Proc Natl Acad Sci U S A **93**(14): 7421-6.
Mumford, D., S. M. Kosslyn, et al. (1987). "Discriminating figure from ground: the role of edge detection and region growing." Proc Natl Acad Sci U S A **84**(20): 7354-8.

Murray, S. O., D. Kersten, et al. (2002). "Shape perception reduces activity in human primary visual cortex." Proc Natl Acad Sci U S A **99**(23): 15164-9.
Nestares, O. and D. J. Heeger (2000). "Robust multiresolution alignment of MRI brain volumes." Magn Reson Med **43**(5): 705-15.
Nothdurft, H. C. (1990). "Texture discrimination by cells in the cat lateral geniculate nucleus." Exp Brain Res **82**(1): 48-66.
Ogawa, S., T. M. Lee, et al. (1990). "Brain magnetic resonance imaging with contrast dependent on blood oxygenation." Proc Natl Acad Sci U S A **87**(24): 9868-72.
Olman, C. A., S. Inati, et al. (2007). "The effect of large veins on spatial localization with GE BOLD at 3 T: Displacement, not blurring." Neuroimage **34**(3): 1126-35.
Paradiso, M. A., S. Blau, et al. (2006). "Lightness, filling-in, and the fundamental role of context in visual perception." Prog Brain Res **155**: 109-23.
Pelli, D. G. (1985). "Uncertainty explains many aspects of visual contrast detection and discrimination." J Opt Soc Am A **2**(9): 1508-32.
Pelli, D. G. (1997). "The VideoToolbox software for visual psychophysics: transforming numbers into movies." Spat Vis **10**(4): 437-42.
Pereverzeva, M. and S. O. Murray (2008). "Neural activity in human V1 correlates with dynamic lightness induction." J Vis **8**(15): 8 1-10.
Press, W. A., A. A. Brewer, et al. (2001). "Visual areas and spatial summation in human visual cortex." Vision Res **41**(10-11): 1321-32.
Purpura, K. P., J. D. Victor, et al. (1994). "Striate cortex extracts higher-order spatial correlations from visual textures." Proc Natl Acad Sci U S A **91**(18): 8482-6.
Ravindranath, B., X Zhang, LR Mujica-Parodi (2007). "Dynamic causal modeling of fMRI time series in response to fearful faces." IEEE 33rd Annual Northeast Bioengineering Conference **1-4244-1033-9**: 41.
Regan, D., and XH Hong (1995). "Two models of the recognition and detection of texture-defined letters compared." Biological Cybernetics **72**(55): 389-396.
Reich, D. S., F. Mechler, et al. (2001). "Temporal coding of contrast in primary visual cortex: when, what, and why." J Neurophysiol **85**(3): 1039-50.
Romani, A., G. Caputo, et al. (1999). "Edge detection and surface 'filling in' as shown by texture visual evoked potentials." Clin Neurophysiol **110**(1): 86-91.
Rosenfield, A., AC Kak (1982). Digital picture processing, 2. New York, Academic Press.
Rovamo, J., O. Luntinen, et al. (1993). "Modelling the dependence of contrast sensitivity on grating area and spatial frequency." Vision Res **33**(18): 2773-88.
Rubenstein, B. S. and D. Sagi (1996). "Preattentive texture segmentation: the role of line terminations, size, and filter wavelength." Percept Psychophys **58**(4): 489-509.
Schofield, A. J. (2000). "What does second-order vision see in an image?" Perception **29**(9): 1071-86.
Schofield, A. J. and M. A. Georgeson (1999). "Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour." Vision Res **39**(16): 2697-716.

Scholte, H. S., J. Jolij, et al. (2008). "Feedforward and recurrent processing in scene segmentation: electroencephalography and functional magnetic resonance imaging." J Cogn Neurosci **20**(11): 2097-109.
Serences, J. T. (2004). "A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI." Neuroimage **21**(4): 1690-700.
Sereno, M. I., A. M. Dale, et al. (1995). "Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging." Science **268**(5212): 889-93.
Shapley, R. M. and J. D. Victor (1978). "The effect of contrast on the transfer properties of cat retinal ganglion cells." J Physiol **285**: 275-98.
Singh, M., JW Jeong, S Kim, W Sungkarat, Y Zhou, TS Kim (2002). "Localization of alpha producing regions in the human brain and their connectivity." Proc Intl Soc Mag Reson Med **10**: 1464.
Skiera, G., D. Petersen, et al. (2000). "Correlates of figure-ground segregation in fMRI." Vision Res **40**(15): 2047-56.
Smith, A. T. and T. Ledgeway (1997). "Separate detection of moving luminance and contrast modulations: fact or artifact?" Vision Res **37**(1): 45-62.
Solomon, J. A. (2000). "Channel selection with non-white-noise masks." J Opt Soc Am A Opt Image Sci Vis **17**(6): 986-93.
Srebro, R. and L. W. Baitch (1991). "Image segmentation and VEP topography." Vision Res **31**(12): 2039-46.
Stephan, K. E., L. Kasper, et al. (2008). "Nonlinear dynamic causal models for fMRI." Neuroimage **42**(2): 649-62.
Sukumar, S. a. S. W. (2007). "Separate first- and second-order processing is supported by spatial summation estimates at the fovea and eccentrically." Vision Res **47**(5): 581-596.
Tootell, R. B., J. B. Reppas, et al. (1995). "Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging." J Neurosci **15**(4): 3215-30.
Turner, M. R. (1986). "Texture discrimination by Gabor functions." Biol Cybern **55**(2-3): 71-82.
Tyler, L. K., E. A. Stamatakis, et al. (2004). "Processing objects at different levels of specificity." J Cogn Neurosci **16**(3): 351-62.
Uludag, K. (2008). "Transient and sustained BOLD responses to sustained visual stimulation." Magn Reson Imaging **26**(7): 863-9.
Vazquez, A. L. and D. C. Noll (1998). "Nonlinear aspects of the BOLD response in functional MRI." Neuroimage **7**(2): 108-18.
Victor, J. D., C. Chubb, et al. (2005). "Interaction of luminance and higher-order statistics in texture discrimination." Vision Res **45**(3): 311-28.
von der Heydt, R., H. Zhou, et al. (2000). "Representation of stereoscopic edges in monkey visual cortex." Vision Res **40**(15): 1955-67.
Voorhees, H. and T. Poggio (1988). "Computing texture boundaries from images." Nature **333**(6171): 364-7.
Wichmann, F. A. and N. J. Hill (2001). "The psychometric function: II. Bootstrap-based confidence intervals and sampling." Percept Psychophys **63**(8): 1314-29.

Wilcox, L. M. and R. F. Hess (1995). "Dmax for stereopsis depends on size, not spatial frequency content." Vision Res **35**(8): 1061-9.
Wilcox, L. M. and R. F. Hess (1997). "Scale selection for second-order (non-linear) stereopsis." Vision Res **37**(21): 2981-92.
Wilcox, L. M. and R. F. Hess (1998). "When stereopsis does not improve with increasing contrast." Vision Res **38**(23): 3671-9.
Yacoub, E., P. F. Van De Moortele, et al. (2005). "Signal and noise characteristics of Hahn SE and GE BOLD fMRI at 7 T in humans." Neuroimage **24**(3): 738-50.
Yesilyurt, B., K Ugurbil, K Uludag (2008). "Dynamics and nonlinearities of the BOLD response at very short stimulus durations." Magn Reson Imaging **26**(7): 853-62.
Zhou, H., H. S. Friedman, et al. (2000). "Coding of border ownership in monkey visual cortex." J Neurosci **20**(17): 6594-611.
Zhu, S., TS Lee, and A Yuille (1995). "Region competition: unifying snakes, region growing and MDL for image segmentation." Proceedings of the Fifth International Conference in Computer Vision: 416-425.
Zipser, K., V. A. Lamme, et al. (1996). "Contextual modulation in primary visual cortex." J Neurosci **16**(22): 7376-89.