University of Nevada, Reno

Neural Coding of Image Blur Assessed by fMRI

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Psychology

by

Katherine Mussell

Dr. Michael A. Webster/Thesis Advisor

May, 2015

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

Katherine Mussell

Entitled

Neural Coding Of Image Blur Assessed By Fmri

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

Michael A. Webster, Advisor

Lars Strother, Committee Member

David Feil-Seifer, Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

May, 2015

Abstract

Blur is a fundamental perceptual attribute of images, but the way in which the visual system encodes this attribute remains poorly understood. We examined the neural correlates of image blur by measuring the fMRI BOLD response to images that varied from focused to either too blurred or too sharpened. Observers viewed grayscale images of natural scenes, filtered by varying the slope of the log amplitude spectra from -1 (strongly blurred) to +1 (strongly sharpened), with RMS contrast equated to the original after filtering. In primary visual cortex (V1) there was higher activation for the in-focus images than for the sharpened or blurred images. Peak responses were similar for focused and sharpened images in foveal V1, while both blurred and sharpened images resulted in lower activity in more peripheral retinotopic regions. Similar patterns were observed in extra-striate areas (V2 to V3), though the differential response to focused images was greatest in V1. These results suggest that focused images provide the strongest neural activity in V1, and run counter to expectations from norm-based or predictive coding in which focus is encoded implicitly as an absence or expected attribute.

Acknowledgements

Completion of this project would not have been possible without the support of Dr. Lars Strother, Dr. David Feil-Seifer, Dr. Gideon Caplovitz, Dr. Ryan Mruczek, Gennadiy Gurariy, and Gennady Erlikhman. Most of all I would like to thank my advisor, Dr. Michael Webster. This research was supported by P20-GM-103650 and EY-10834.

Contents

List of Figures

Neural Coding of Image Blur Assessed by fMRI

Blur is an important dimension of normal vision, and is a feature that the visual system is highly sensitive to (e.g. in judging image quality or the relative depth of objects (Brady & Field, 2000; Field, 1987; Mather & Smith, 2002; Olshausen, 1996)), and constantly adjusting for (e.g. through processes such as accommodation of the eye's optics or neural adaptations (von Helmholtz & Southall, 1924)). Yet despite its salience as visual feature, the stimulus cues defining blur, and how these are represented within the visual system, are poorly understood. Here we examined the nature of the neural representation of blur, by using functional magnetic resonance imaging (fMRI) to compare neural activity to images that were in focus or either blurred or sharpened.

Blur is generally associated with a loss in the fine detail of images or more gradual changes in the profiles of edges (Field & Brady, 1997; Webster, Georgeson, & Webster, 2002). Most natural images have a characteristic distribution of energy across different spatial scales that is captured by the image power or amplitude spectrum. In particular, the amplitude spectra of natural images tend to fall inversely with increasing spatial frequency, or as $1/f$, so that on a log-log scale the relationship between amplitude and spatial frequency follows a slope of -1. It has been repeatedly suggested that visual coding is well matched to this structure (Field, 1987; Olshausen, 1996; Webster & Miyahara, 1997). For example, the spatial frequency bandwidths of cortical simple cells increases with spatial frequency (e.g. roughly as $f/1$), so that the responses of the cortex to a $1/f$ spectrum remain roughly constant. Thus this suggests that the visual system in fact gives equal "neural" weight to different spatial scales, even though these are

disproportionately weighted toward lower frequencies in the physical image. A potential signature of image focus might thus be equal neural responses to different frequency ranges in the image, though Field and Brady (Brady & Field, 2000) have shown that it is not only the total average amplitude spectrum but rather the spectrum corrected for the density of the spatial structure (since images that have fewer edges will have steeper spectra even though the spectra of the individual edges are $1/f$).

Most accounts of image blur are based on optical blur, which necessarily degrades the retinal image and thus removes finer details from the image. This in turn steepens the amplitude spectrum. By this account a focused image is always the sharpest image that can be achieved, and is at the end of the continuum of possible blur. However, it is also possible to create an image that is too sharp by reducing the contrast at lower rather than higher frequencies in the amplitude spectrum, thus creating an image in which the spectrum has a slope shallower than $1/f$. By this account blur is instead at the middle of the continuum of possible stimulus levels. While over-sharpening cannot be achieved optically, it could occur neurally if the sensitivity of high spatial-frequency cells is too great (or low-frequency cells too weak). In terms of these neural responses, focus thus might correspond to a unique neural signature, when the responses across spatial frequency are constant and thus the neural response is "unbiased." This unbiased response further represents a unique norm or prediction about the expected spatial structure of the world.

In this study, our aim was to test for a neural correlate of a norm in the visual coding of blur, and in particular to assess whether overall neural activity might be

weakest for focused images. To examine this, we filtered the slope of the amplitude spectra of natural images to create images that varied from “too blurry” to “too sharp.” We then used fMRI in order to measure an analog of neural activity known as blood oxygen level dependent (BOLD) signal.

There are several reasons to expect that the BOLD signal would increase as the stimuli move away from the in-focus images in either direction on the blur spectrum. The first is that the attribute of blur could be encoded by a norm-based representation, in which stimulus levels are represented according to how they differ from an expected level (the norm). This relative coding scheme may underlie the way in which many visual features are encoded, and predicts that neural responses will increase as stimuli differ further from the norm. Some evidence for this coding scheme has been shown previously for fMRI responses. For example, Loffler et al. presented faces that varied in how they diffed relative to a geometrically determined average face. It was found that activity for the “norm” or average face was lower than activity for faces that were more distinctive (Loffler, Yourganov, Wilkinson, & Wilson, 2005). A similar norm-based coding pattern has been found for color, in which each color may be defined by how (and how much) it differs from a neutral gray. Area V4 has the lowest activation associated with achromatic stimuli, and increasing activation as more color is added (Wade, Augath, Logothetis, & Wandell, 2008).

A norm-based code for blur has been suggested by studies of perceptual adaptation (Webster et al., 2002). Specifically, adapting to a blurred or sharpened image biases the image that appears in focus, and thus may reflect a “renormalization” of blur perception. If this norm were encoded in a similar way to gray or to an average face, then

it would be expected that the lowest BOLD signal would be associated with in-focus images (the norm), and BOLD signals would increase with blurred and sharpened images.

Norms are closely related to ideas of predictive coding, in which the visual system builds an expectation or prediction about the world so that it only needs to represent the errors or deviations from these predictions. Some studies have found that the amount that the stimulus differs from our expectations relates directly to an increase in the BOLD signal (Hesselmann, Sadaghiani, Friston, & Kleinschmidt, 2010; Summerfield et al., 2006). Thus predictive coding would again predict that we would have the lowest neural activity for in-focus images, as these are the most common or expected image type we encounter.

The final reason to expect that in-focus images would result in less neural activity is based on the literature of visual discomfort. In a study done by Juricevic et al. (Juricevic, Land, Wilkins, & Webster, 2010) participants were asked to rate their level of visual discomfort while observing stimuli that had been manipulated in various ways. Visual discomfort increased as stimuli deviated from the $1/f$ spectra of natural scenes. Previous studies have shown that discomfort may result from higher levels of neural activity (Fernandez & Wilkins, 2008; Haigh et al., 2013; O'Hare & Hibbard, 2011). Therefore, the fact that blurred or sharpened images are perceived as more uncomfortable could indicate that they elicit greater activity.

The current study was designed to directly probe for a neural norm in the visual encoding of blur, by measuring BOLD responses to sets of images that were identical except for the slope of their amplitude spectra. This allowed us to assess whether

variations in image blur could be tracked with fMRI, and if so how these variations impacted responses in early stages of the cortical visual system.

Methods

Participants

Six observers participated in the experiment (1 female). These included author KT and 5 graduate students who were unaware of the specific aims of the study. One of the latter's data was excluded due to an inability to properly align the EPI (echo planar imaging) with the anatomical images (interference from dental brace). All observers had normal or corrected-to-normal vision and completed a standard safety screening. Participation was with informed consent and followed protocols approved by the university's IRB.

Design

A rapid-event related design was utilized with a display duration of 2 sec and a variable ISI of 2 to 6 sec (Figure 1). Participants were run in either one or two sessions. Each type of stimulus was presented 10 times within each run, and each run lasted 340 sec. Stimulus types were randomized using a counterbalanced sequence. The task was a simple fixation task wherein participants were asked to press a button when the fixation point changed from red to green. This was utilized to ensure that participants maintained fixation and were attentive throughout the entire scanning session.

Stimuli

Stimuli consisted of 18 achromatic images of natural scenes selected from the McGill Calibrated Colour Image database (Olmos, 2011) and converted to grayscale. The depicted scenes included far landscapes, and forest scenes. The images were filtered by

varying the slope of the log amplitude spectra from -1 (strongly blurred) to +1 (strongly sharpened) relative to the original slope (Webster et al., 2002). Mean luminance and RMS contrast were equated to the original after filtering. Each image was 256x256 pixel (2.71^o x 2.71^o visual angle), and tiled to fill the entire screen (18.74^o x 11.96^o visual angle) with an image centered behind the fixation point. The stimuli were presented using the Psychophysics Toolbox (Brainard, 1997) for MATLAB (Mathworks Inc., Natick, MA). Stimuli were presented on a Cambridge Research System (Kent, UK) LCD BOLD screen display (60-Hz refresh rate) outside of the scanner room, and were viewed through a mirror attached to the head coil.

fMRI Data Acquisition

Scans were collected at the University of California, Davis Skyra facility using a 3-T Siemens MR system (Siemens Healthcare, Erlangen, Germany) and 32-channel head coil. T2* weighted images were collected sequentially in an anterior to posterior direction (32 total slices with nearly whole-brain coverage, matrix = 80x80, voxel size = 3x3x3 mm3, TR = 2000ms, TE = 30 ms, FA = 71 deg.).

Retinotopic mapping as well as localizer scans were performed prior to the study in a separate scanning session. T1 images were collected prior to functional images (voxel size = 2x2x2, MPRAGE: 208 sagittal slices, $0.9mm^2$ in-plane voxel resolution, matrix size = 256 x 256, slice thickness = 0.95 mm, FOV = 243 x 243 x 187 mm, TE = 4.33 ms, TR = 10 ms, flip angle = 7^o). Along with structural images, functional localizers were collected in order to create individualized regions of interest for each participant. The retinotopic mapping stimuli used were flickering checkerboards (4 Hz) varying in eccentricity and then in polar angle (Arcaro, McMains, Singer, & Kastner, 2009;

Swisher, Halko, Merabet, McMains, & Somers, 2007). The moving polar angle wedge covered 45^{o} and spanned from 0.5^{o} to 13.5^{o} eccentricity for half the runs, and 8^{o} to 13.5^{o} for the other half. For the eccentricity mapping runs, the stimulus was a moving ring which cycled between 0^{o} and 13.5^{o} from fixation. Borders between topographic areas were defined by reversals in polar angle representations. Topographic regions were defined using AFNI (http://afni.nimh.nih.gov/afni/Cox, 1996) and SUMA (http://afni.nimh.nih.gov/afni/suma) (Saad, Reynolds, Argall, Japee, & Cox, 2004).

Data Analysis

All data were analyzed using AFNI (http://afni.nimh.nih.gov/afni/Cox, 1996) (Saad et al., 2004) including time slice correction and motion correction (checked by hand and realigned using the run with the least motion correction). EPI and anatomy alignment were also checked by hand, and functional volumes were smoothed using 6 x 6 Gaussian kernels. A whole-brain general linear model was used to highlight all regions in which activity was greater than baseline. An ROI analysis then separated activity by brain region. For analyses V1 was further divided into two sub-regions based on eccentricity mapping, corresponding to “foveal” and “non-foveal” visual areas (Figure 2). ROI’s were drawn using SUMA (http://afni.nimh.nih.gov/afni/suma) (Saad et al., 2004).

Results

A whole-brain GLM revealed that activity greater than baseline was primarily limited to early visual areas. Therefore, we restricted our analysis to these regions. Figure 3 highlights whole brain contrasts for three representative subjects. The top row reveals all regions in which activity for in-focus images was greater than activity for blurred images. Row 2 reveals all regions in which activity was greater for in-focus images than

for sharpened images. Anecdotally, it also seemed that the contrasts comparing in-focus to blur revealed greater activity toward the more polar portion of occipital cortex, corresponding to the central visual field, while contrasts comparing in-focus and sharpened images revealed activity shifted forward (Figure 3). This motivated our decision to divide early visual areas into foveal and peripheral sub-regions.

For all ROI analyses, we averaged beta weights across the two hemispheres, as the GLM revealed very symmetric patterns of activation. For region V1, the in-focus images produced the highest level of activation. The effects of blur were confirmed with a one-way repeated measures ANOVA ($F(4,16)=10.837$, $p<0.001$). Further post-hoc comparisons showed that the in-focus images led to greater activity than both the most-blurred images ($q(4)=8.027$, $P<0.001$) and the sharpest images ($q(4)=6.961$,$P<0.001$). In contrast, in areas V2 and V3 the activation levels did not significantly differ for the different levels of filtering (Figure 4).

For further analysis we divided V1 into two sub-regions based on eccentricity mapping (Figure 2). These results revealed that in the foveal region, there was a significant difference between in-focus images and blurred images ($q(4)=13.56$, $P<0.001$), but no difference between in-focus images and sharpened images ($q(4)=3.229$, $P=0.088$). In the non-foveal region of V1, there was a significant difference between the in-focus images in both the blurred ($q(4)=9.750$, $P<0.001$) and sharpened ($q(4)=10.799$, $P<0.001$) directions (Figure 5). Again, for V2 and V3 the activation levels in both foveal and peripheral locations did not significantly vary with the blurring (Figure 6).

Discussion

To summarize, our results revealed that in-focus images elicited the strongest BOLD response in early visual areas. Moreover, the effects of image slope changes on the BOLD responses were most apparent in V1 and weaker in higher visual areas. We consider these two results in turn.

Previous studies have also reported a decrease in the BOLD response with optical blur. Mirzajani et al. (Mirzajani, Sarlaki, Kharazi, & Tavan, 2011) used positive lenses to induce myopia in participants with otherwise normal vision. Participants were then presented with sinusoidal gratings while neural activity was measured using fMRI and EEG. Mirzajani et al. (Mirzajani et al., 2011) found decreased activity for stimuli that had been blurred through the lenses, and higher activity for in-focus gratings. However, optical blur reduces the overall contrast of the images and thus the changes in BOLD responses are likely to reflect these contrast losses (Boynton, Demb, Glover, & Heeger, 1999). In our study we attempted to circumvent this problem by equating images for a constant RMS contrast after filtering. While there is no well-defined metric for defining or equating contrast in complex images, our RMS manipulation kept the total power constant. Moreover, the fact that the BOLD responses did not vary monotonically with the image slopes suggest that they cannot be accounted for by changes in contrast.

The fact that BOLD responses reached a peak rather than a minimum for the focused images appears at odds with simple predictions about predictive or norm-based coding. Again if focus were represented only implicitly, as an absence of blur, then we might have expected the focused images to lead to the lowest activation levels. However, our results do not necessarily preclude a norm-based representation. In particular, our

measurements were based on global responses to images with complex properties. Thus the stimuli included many attributes (textures, objects, etc.) that were not normalized. If the code for blur reflected responses only within a subset of the neural population, then these may have been masked by the overall activity to the many other patterns and features in the images.

However, why should focused images then lead to greater activity? A number of factors may be at play. For example, the cortical response to contrast is known to exhibit a compressive nonlinearity (Boynton et al., 1999; Kwon, Legge, Fang, Cheong, & He, 2009), with a typical functional form expressed by:

$$R = R_{max} \frac{C^{n+m}}{(C^n + C_{50}^n)}$$

Because of this, the trade-offs we introduced between adding more amplitude at some scales while subtracting from others may not have led to equivalent responses (e.g. because the response to excess contrast may have saturated). By this account the focused images may have given the largest responses because they kept the most channels in the middle of their response range, thus avoiding saturation.

A second nonlinearity in the cortical coding that might have impacted the responses is contrast gain control, in which the responses of each neuron are modulated by the pooled responses of the cells (Carandini & Heeger, 2012). This normalization can result in a "winner take all" response when the stimulus is biased (Busse, Wade, & Carandini, 2009), and thus could suppress the overall response when the stimulus was blurred or sharpened compared to the balanced focused stimulus.

Finally, while blurring or sharpening introduces an attribute to the image, it also removes attributes, by biasing the stimulus to appear composed primarily only of fuzzy structure or fine detail. In contrast, $1/f$ or focused images appear to have structure at multiple scales that appear equally salient (Field & Brady, 1997). Thus varying the slope away from $1/f$ might actually have reduced the heterogeneity of structure in the image, potentially again reducing the total neural response.

Whatever the basis for these effects, it is also notable that the differential responses to blur were largely lost in higher visual areas. Higher areas tend to show greater stimulus invariance including less dependence on stimulus contrast (Avidan et al., 2002). Our results are consistent with the idea that the relative contrast at different spatial scales that is the characteristic feature of blurred or sharpened images may be represented at earlier levels like V1 where better contrast information is retained.

References

Arcaro, M. J., McMains, S. A., Singer, B. D., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *The Journal of Neuroscience, 29*(34), 10638-10652.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., & Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *Journal of Neurophysiology, 87*(6), 3102-3116.

Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). Neuronal basis of contrast discrimination. *Vision research, 39*(2), 257-269.

Brady, N., & Field, D. J. (2000). Local contrast in natural images: normalisation and coding efficiency. *Perception-london-, 29*(9), 1041-1056.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*, 433-436.

Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of concurrent stimuli by population activity in visual cortex. *Neuron, 64*(6), 931-942.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*(1), 51-62.

Fernandez, D., & Wilkins, A. J. (2008). Uncomfortable images in art and nature. *Perception, 37*(7), 1098.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Josa a, 4*(12), 2379-2394.

Field, D. J., & Brady, N. (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision research, 37*(23), 3367-3383.

Haigh, S. M., Barningham, L., Berntsen, M., Coutts, L. V., Hobbs, E. S., Irabor, J., et al. (2013). Discomfort and the cortical haemodynamic response to coloured gratings. *Vision research, 89*, 47-53.

Hesselmann, G., Sadaghiani, S., Friston, K. J., & Kleinschmidt, A. (2010). Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PLoS One, 5*(3), e9926.

Juricevic, I., Land, L., Wilkins, A., & Webster, M. A. (2010). Visual discomfort and natural image statistics. *Perception, 39*(7), 884.

Kwon, M., Legge, G. E., Fang, F., Cheong, A. M. Y., & He, S. (2009). Adaptive changes in visual cortex following prolonged contrast reduction. *Journal of vision, 9*(2), 20.

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature neuroscience, 8*(10), 1386-1391.

Mather, G., & Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception-london-, 31*(10), 1211-1220.

Mirzajani, A., Sarlaki, E., Kharazi, H. H., & Tavan, M. (2011). Effect of lens-induced myopia on visual cortex activity: A functional MR imaging study. *American Journal of Neuroradiology, 32*(8), 1426-1429.

Olmos, A. (2011). Kingdom FAA (2004) McGill calibrated colour image database. *URL http://tabby. vision. mcgill. ca. Last accessed*, 05-08.

Olshausen, B. A. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*(6583), 607-609.

O'Hare, L., & Hibbard, P. B. (2011). Spatial frequency and visual discomfort. *Vision research, 51*(15), 1767-1777.

Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., & Cox, R. W. (2004). *SUMA: an interface for surface-based intra-and inter-subject analysis with AFNI.* Paper presented at the Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science, 314*(5803), 1311-1314.

Swisher, J. D., Halko, M. A., Merabet, L. B., McMains, S. A., & Somers, D. C. (2007). Visual topography of human intraparietal sulcus. *The Journal of neuroscience, 27*(20), 5326-5337.

von Helmholtz, H., & Southall, J. P. C. (1924). Mechanism of accommodation.

Wade, A., Augath, M., Logothetis, N., & Wandell, B. (2008). fMRI measurements of color in macaque and human. *Journal of Vision, 8*(10), 6.

Webster, M. A., Georgeson, M. A., & Webster, S. M. (2002). Neural adjustments to image blur. *Nature neuroscience, 5*(9), 839-840.

Webster, M. A., & Miyahara, E. (1997). Contrast adaptation and the spatial structure of natural images. *Josa a, 14*(9), 2355-2366.

Figures

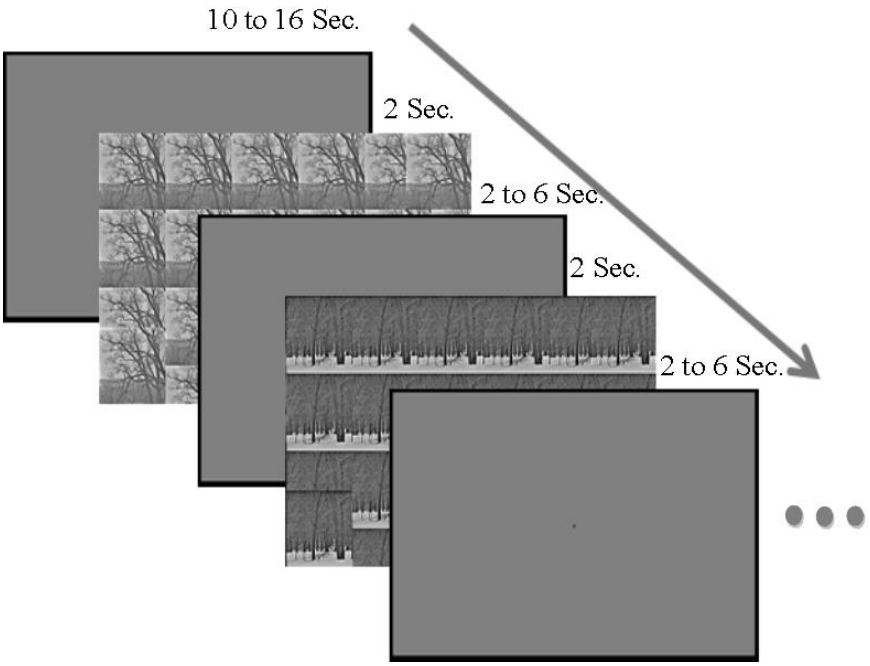

Figure 1. Experimental design: rapid event related design with 2 sec. display and 2 to 6 sec. jittered ISI.

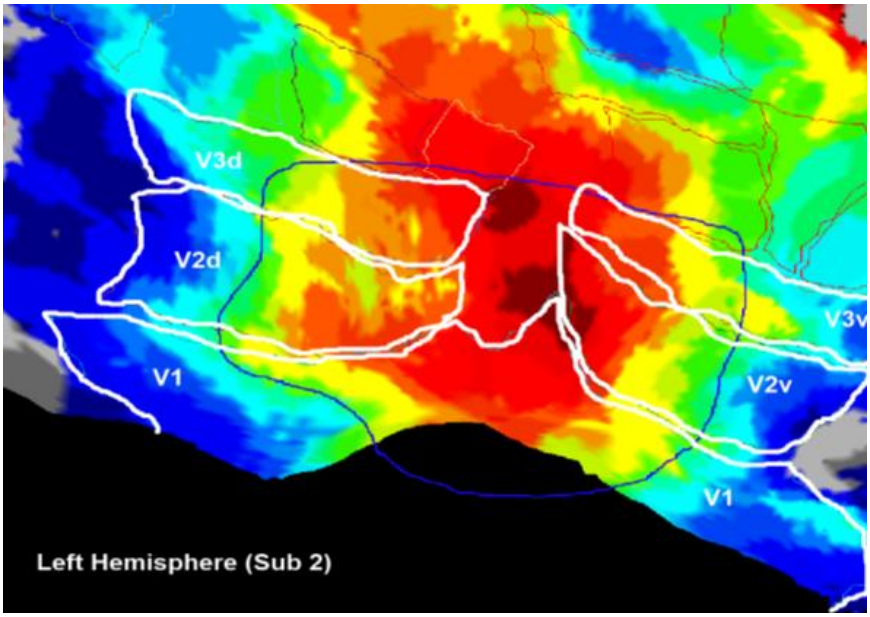

Figure 2. Example of foveal and peripheral sub-regions. Areas V1, V2, and V3 were each divided into two sub-regions using independent eccentricity mapping. The activation for the eccentricity stimuli overlaid on a flat surface map.

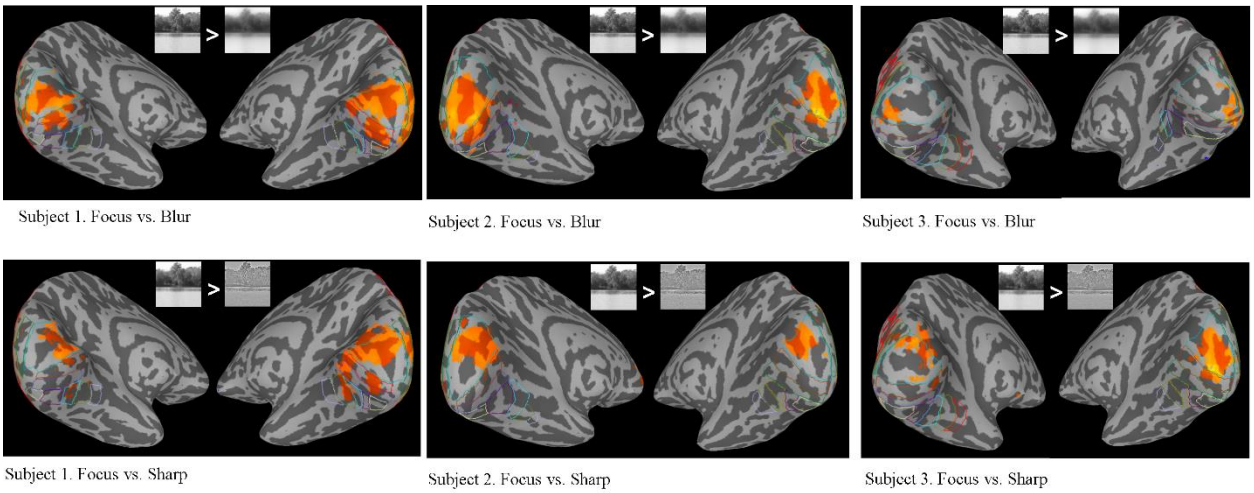

Figure 3. Results: Whole brain contrasts. The whole brain contrasts for 3 representative subjects. The highlighted regions represent positive activity. The top row reveals areas where in focus activity was greater than blurred. The bottom row reveals areas where in focus activity was greater than sharpened.

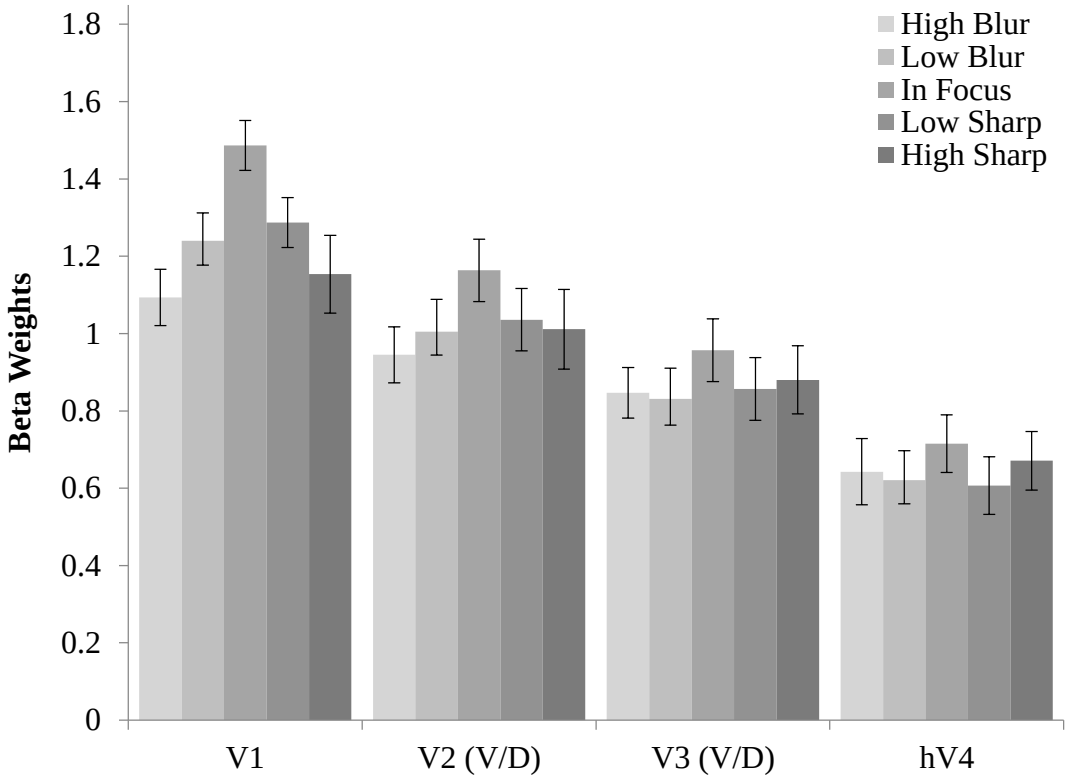

Figure 4. Results: ROI analysis for areas V1, V2, V3, and hV4. Bars represent the standard error. Beta weights calculated by using a general linear model. (Key: Lightest to Darkest – High Blur, Low Blur, In Focus, Low Sharp, High Sharp)

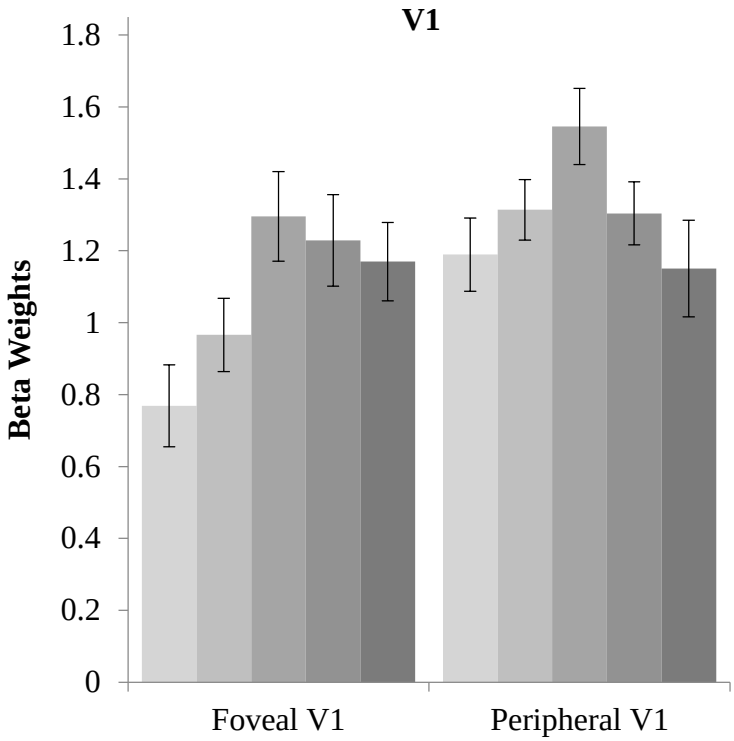

Figure 5. Results: ROI analysis for foveal and peripheral sub-regions of V1. Eccentricity based sub-regions of V1. Bars represent standard error. Beta weights calculated using a general linear model. (Key: Lightest to Darkest – High Blur, Low Blur, In Focus, Low Sharp, High Sharp)

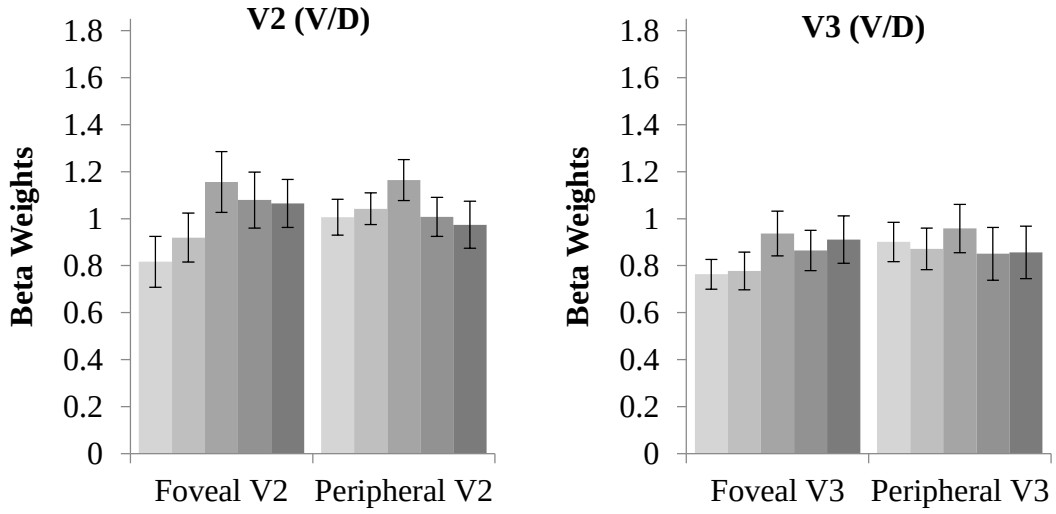

Figure 6. Resutls: ROI analysis for foveal and peripheral sub-regions of V2 and V3. Eccentricity based sub-regions for V2 and V3. Bars represent standard error. Beta weights calculated using a general linear model. (Key: Lightest to Darkest – High Blur, Low Blur, In Focus, Low Sharp, High Sharp)