University of Nevada, Reno

Color perception in anomalous trichromats:
Neuroimaging investigations of neural compensation for losses in spectral sensitivity

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Psychology

by

Katherine E.M. Tregillus

Dr. Michael A. Webster/Dissertation Advisor

August, 2017

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

KATHERINE E.M. TREGILLUS

Entitled

Color perception in anomalous trichromats: Neuroimaging investigations of neural compensation for losses in spectral sensitivity

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

DOCTOR OF PHILOSOPHY

Michael A. Webster, Advisor

Michael A. Crognale, Committee Member

Fang Jiang, Committee Member

Dennis Mathew, Committee Member

Alexander van der Linden, Graduate School Representative

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

August, 2017

Abstract

Anomalous trichromats have reduced sensitivity to the L-M dimension of color space due to the reduced separation between the spectral sensitivities of their L and M cones. Despite this, previous work suggests that these observers may perceive the world to be much more colorful than their cone sensitivities would predict, potentially because of long-term adaptation that amplifies the weakened chromatic signals provided by the cones. Most of the evidence for this gain adjustment rests on subjective measures of color appearance or color salience. In the present study, we tested for neural correlates of color compensation by using fMRI to compare the cortical responses to chromatic stimuli in normal and anomalous observers. Thresholds were collected for a total of 7 anomalous trichromats (3 deuteranomals and 4 protanomals), and 6 color normal controls. Initial results showed that chromatic thresholds for the L-M axis did not predict BOLD responses, indicating neural compensation in early visual areas. In an additional experiment, we used an attentionally demanding task to ensure that top-down influences were limited. We also collected retinotopic mapping in order to independently define early visual areas (V1, V2, V3, and hV4). In this case, the group-averaged BOLD responses to L-M stimuli were not significantly greater than responses predicted by threshold, but individual participants did show evidence of compensation. The same was true when responses were normalized to responses to S-axis stimuli. Our results thus provide evidence for compensatory amplification, but suggest that the degree of compensation varies across individuals.

Acknowledgements

I couldn't have completed this project without the help and support of many people. Thank you to all of my participants. Thank you to my lab mates Kara Emery, Siddart Srivatsav, and Ivana Illic. Thank you to my collaborators, Michael Crognale, Stephen Engel, Ichiro Kuriki, and Donald Macleod. Thank you to members of my dissertation committee, Fang Jiang, Dennis Matthew, and Alexander Van der Linden. Thank you to those who offered help with fMRI and facilities, Lars Strother, Zhiheng Zhou, Sean O'Neil, and Larry Messier. Thank you to John Erik Vanston, for his collaboration and support. Thank you to my partner, Samuel Tregillus for all of his encouragement and patience throughout my entire time in graduate school, but especially through the dissertation process. And finally, thank you to my advisor, Michael Webster. I really could not have asked for a more fantastic mentor, or a better graduate experience. Financial support and equipment were provided by NIH P20GM103650 and NIH EY-10834, Renown Hospital in Reno, and the University of Nevada, Reno.

Contents

Abstract i

Acknowledgements ii

List of Figures v

Introduction 1

Methods 6

Anomaloscope 6

Participants 7

Stimuli 7

Threshold Procedure 8

fMRI Procedure 9

Fixation Tasks 9

fMRI Data Preprocessing 10

Retinotopic Mapping 10

Results 12

Thresholds 12

Experiment 1: Simple Fixation 12

Fitting Contrast Response Function 13

Statistical Comparisons 14

Experiment 2: High-Attentional Fixation 14

Comparisons across Experiments 1 and 2 15

Fitting CRF 16

Statistical Comparisons within ROIs 16

ROIs as Within Subject Comparison 18

Comparisons of S/(L+M) Responses 18

Data Normalized to S/(L+M) Responses 19

Statistical Comparisons within ROIs 19

ROIs as Within Subject Comparison 20

Comparisons of Scaling Factors. 21

Basis for changes in the CRF 22

Individual Observers 23

Dichromatic Observer 24

Discussion 25

References 31

Figures 38

List of Figures

Figure 1. Example of shifted cones in anomalous trichromacy, from (M. Neitz & Neitz, 2000). A) The sensitivity spectra for each of the three cone types in a trichromat. B & C) The three cone types present in deuteranomalous and protanomalous individuals. 38

Figure 2. A) An original color image. B) The image as seen by a protanomalous observer assuming adaptation only in the cones to match the average spectrum. C) The image assuming contrasts are rescaled by post-receptoral gain. Figure from (Webster et al., 2010). 38

Figure 3. fMRI stimuli, top two rows are the L/L(+M) gratings and the bottom two are the S/(L+M) gratings. Each was presented at 4 contrast levels, 80, 40, 20, and 10 in units of nominal multiples of threshold. These radial sine wave gratings were also used in collecting detection thresholds. 39

Figure 4. An example of ROIs for 1 observer (C-02) on an inflated brain. 40

Figure 5. Raw threshold results for the two chromatic axes, represented in nominal multiples of threshold. The x-axis is subject number, with control participants in gray. . 40

Figure 6. Whole brain GLM (contrast: all conditions greater than fixation), for two representative observers. C-03, control observer. DA-04, deuteranomalous observer. ... 41

Figure 7. fMRI results from Exp. 1, Simple Fixation task. Shown here is all activity greater than baseline using a GLM from an activation-based ROI in early visual cortex, and averaged across observers. Error bars represent the standard error of the mean. 41

Figure 8. Solid lines are the BOLD responses from Exp. 1, simple fixation task. The dashed lines are CRFs fitted to the data. The black, dashed line is the CRF of the control

data, scaled to the threshold ratio of control to AT. Red are the L/(L+M) data, and blue are the S/(L+M) data. 42

Figure 9. The mean responses at each contrast level in percent signal change for Exp. 2, the high-attentional fixation task. Each row is activation in a region of interest, combined across the left and right hemispheres (V1, V2v, V3v, and V4), column 1 shows response to L/(L+M) stimuli, and column 2 shows activation to S/(L+M) stimuli. 44

Figure 10. The BOLD response to L-M stimuli for AT observers and controls, from Experiment 1 (low attention) and Experiment 2 (high attention) to L/(L+M) stimuli. Only observers who participated in both experiments are included here. 46

Figure 11. Each plot shows the activation for a region of interest (V1-V4) for the L vs. M stimuli. The solid lines are the data, and the dashed lines are the fitted CRF functions. The gray is the control data, the red is the anomalous trichromat data, both averaged across observers. The black, dashed function is the CRF fitted to the control data, scaled using the threshold ratio of anomalous trichromats to controls. 47

Figure 12. Individual plots for each AT observer (row), and ROI (column). The AT observer's data is in red, and the control data is in gray. Solid lines represent the BOLD respones and dashed lines are the best fit CRF. The black line is the threshold prediction for each AT observer. 48

Figure 13. Responses for each individual and ROI plotted against threshold predictions derived from the threshold ratio. Data falling above the solid line indicate responses where data was greater than predicted by the thresholds. 49

Figure 14. Each plot shows the activation for a region of interest (V1-V4) for the S/(L+M) stimluli. The solid lines are the BOLD repsonses averaged across participants, and the dashed lines are the fitted CRF functions. The gray is the control data, the blue is the anomalous trichromat data, both average across observers. 50

Figure 15. Each plot shows results from a region of interest analysis (V1-V4) for the L/(L+M) data, normalized to the S/(L+M) data, where the response to S/(L+M) at contrast 80 is 1 and the baseline is 0. The solid lines are the actual data averaged across participants, and the dotted lines are the best fit CRF (green is AT, and gray is control). The black line was derived using the constants from the control data, scaled by the threshold ratio of AT to control. 51

Figure 16. The individual data, renormalized to each participants' S/(L+M), contrast 80 response. Rows are the individual data for each AT participant, columns are the regions of interest (V1-V4). "DA" indicates deuteranomalous and "PA" indicates protanomalous. 52

Figure 17. Threshold ratios (L/(L+M) vs. S/(L+M)) and scaled MRI activation ratios (L/(L+M) vs. S/(L+M)) for individual observers (x-axis). A ratio of 1 would indicate equal responses for the L-M and S axes. 53

Figure 18. SS error of the scaled fits from controls to AT observers, vs. the scaled fits derived from modulating either rMax (red) or c50 (blue). 54

Figure 19. Activation to chromatic contrast levels for one protanopic participant, as well as the mean activation across AT observers for areas V1, V2v, V3v, and V4. Absent bars indicate no significant activation above baseline. As in earlier graphs, the left column

shows activation to the L/(L+M) stimuli, and the right column shows activation to the S/(L+M) stimuli. 56

Figure 20. Results from Vanston et al., 2017. Suprathreshold contrast matches were consistent with predictions based on AT thresholds. 57

Color perception in anomalous trichromats: Neuroimaging investigations of neural compensation for losses in spectral sensitivity

A central issue in sensory neuroscience and perceptual psychology is how sensory systems are calibrated to represent information. Optical and neural properties of the visual system vary dramatically between observers, or within the same observer at different locations (e.g. central vs. peripheral viewing) or times (e.g. during development or aging). Yet observers tend to agree on their perceptual experiences more than their sensitivity differences might predict. For example, the lens of older observers is much yellower and thus screens the light reaching the retina very differently, yet the stimulus that appears white remains very stable as we age (Werner and Schefrin, 1993; Delahunt et al., 2004). Such results suggest that visual processes and representations are calibrated to compensate for the inherent sensitivity limits of the observer. For instance, the perception of white may correspond to the average spectrum we are exposed to, and thus remain similar for two observers exposed to the same world, even if their eyes filter the spectrum in different ways (Webster and Leonard, 2008). In this project my aim is to explore the nature and extent of these compensatory processes in observers with markedly different color sensitivity, by testing individuals with color deficiencies.

Typical color perception in humans is trichromatic because it depends on sensing the light with three different types of cone photoreceptors (known as long-, middle-, and short-wave receptors (L, M, and S) according to the wavelengths they are most sensitive to). The signals in the cones are then compared in color-opponent channels that contrast the cone signals, e.g. to detect whether the L or M cone response was stronger. However,

in approximately 7% of the Caucasian male population, the L or M cone is affected, creating a color deficiency (MacLeod and Hayhoe, 1974; Neitz and Neitz, 2000; Neitz and Neitz, 2011). This can lead to a complete loss of the cone type (dichromacy) or a change in spectral sensitivity such that the peak is shifted to near that of the other cone (anomalous trichromacy). Examples of the spectral sensitivities for anomalous trichromats are shown in Figure 1. The reduced separation between the cones leads to a smaller "difference" signal and thus reduced sensitivity to color (Nathans et al., 1986; Merbs and Nathans, 1992b, a; Asenjo et al., 1994).

This reduction in sensitivity results in reduced performance in color vision tasks that should be well predicted by the spectral properties of the cones, yet the extent of individual deficits seems to vary greatly. Anomalous trichromats' performance on color discrimination tasks can range from essentially dichromatic to nearing performance of trichromats (Deeb et al., 1992; Neitz et al., 1996; Sanocki et al., 1997; Crognale et al., 1998; Shevell et al., 1998). Models attempting to explain the variation seen in anomalous trichromats have failed to account for all variations using genotype (Barbur et al., 2008), creating a need to find objective measures of anomalous trichromatic color responses. Crognale et al. have previously shown that chromatic visual evoked potentials acquired from participants with color vision deficiencies provided reasonable diagnostic criteria (Crognale et al., 1993), with more severe color deficiencies resulting in higher latencies, including differences among AT observers. More recently, Rabin et al. showed that cone-specific VEPs produce accurate diagnoses of color vision deficiencies, including anomalous trichromats (Rabin et al., 2016).

Currently, it remains unclear to what extent the world "looks" less colorful to anomalous trichromats. Previous work investigating the color responses of anomalous trichromats at a perceptual level have indicated that it may be very different than normal trichromats. For example, Müller et al. (Müller et al., 1991) used reaction times as a measure of color discriminability. Deuteranomalous observers were asked to judge whether two colors were the same or different, and the response times were used to scale the perceptual differences using a multidimensional scaling procedure. They found that the anomalous trichromats' reconstructed color spaces were markedly different from color normal observers. Similarly, Regan and Mollon (Regan and Mollon, 1997) showed in a perceptual grouping task that the relative salience of stimuli varying along the affected color axis was reduced compared to color normal observers.

Despite these deficits, many have theorized that anomalous trichromats may still have a richer color experience. First, the difference signals between the cones are very asymmetric even in normal trichromats. Specifically, the L and M cones are normally separated by only 30 nm compared to more than 100 nm between the peaks of the M and S cones. Thus if color percepts were determined by the magnitude of the cone inputs to the color opponent mechanisms then all observers should experience reds and greens as less salient. Yet perceptually the world appears to vary as much in red and green as blue and yellow, or bright and dark (McDermott and Webster, 2012). Second, in normal trichromats chromatic sensitivity is higher for L vs. M than S vs. LM signals or luminance signals, as if compensated for the reduction in the inherent differences (Chaparro et al., 1993; von der Twer and MacLeod, 2001). That is, the L vs. M signals, while inherently weaker, may be amplified in the post-receptoral channels. There is also a

strong theoretical rationale for these gain adjustments. Each neuron has a limited dynamic range of responses, and neurons are thought to be tuned to adjust this range so that it optimally represents the range of available inputs (von der Twer and MacLeod, 2001; Webster, 2011).

There are intriguing suggestions that these neural compensations also occur in anomalous trichromacy. Though Regan and Mollon found reduced salience to L-M stimuli in most anomalous observers, for some the L-M salience was nevertheless stronger than predicted by their cone sensitivities. More recently, Boehm et al. (Boehm et al., 2014) used multi-dimensional scaling to compare the perceptual distances between different colors for normal or color deficient observers. Performance on color screening tests predicted that the anomalous trichromats should rate colors varying along the L-M axis to be on average 4 times less distinct. However, the subjective judgments for the two groups instead appeared similar, and thus again very different from the settings predicted by their cone sensitivities.

One mechanism that could account for these compensatory adjustments is sensory adaptation, in which neural signals adjust to the average and range of the ambient stimulation (Webster et al., 2010). If the contrast is too low, then a neuron may increase its sensitivity so that the range of outputs is maintained (MacLeod, 2002; Rieke and Rudd, 2009). Whether this affects sensitivity depends on whether the adaptation occurs before or after the sites at which noise limits performance. If the adaptation occurs before the noise, then it could in principle amplify only the signal, effectively discounting the cone sensitivity losses (Figure 2).

Alternatively, an adaptive adjustment at later stages might amplify both signal and noise, leading to increased responses to red and green even if the sensitivity to these colors remains poor. Both types of effects have been observed in anomalous trichromacy (MacLeod, 2002).

However, most of the current evidence for surpathreshold color compensation in anomalous trichromats rests on subjective judgments of appearance (e.g. how similar two colors appear or their relative salience). Such judgments cannot readily discriminate between actual neural response gains vs. criterion effects (e.g. observers may simply be judging the colors relative to the range they are accustomed to seeing). Our goal was to evaluate color compensation by instead using objective measures of neural gains.

fMRI has already proven to be a useful measure of chromatic responses and adaptive gain. Previous Chromatic responses are well characterized across different regions, and different chromatic angles have been shown to elicit distinct BOLD responses (Engel et al., 1997; Mullen et al., 2007; Wade et al., 2008; Kuriki et al., 2015; Mullen et al., 2015; Johnson and Mullen, 2016). fMRI has also been used to reveal potential long-term chromatic adaptation effects to the colors that are statistically more likely to appear in the natural environment, i.e. those colors that fall along the daylight axis of color space. (Goddard et al., 2010; Stringham et al., 2013; Mullen et al., 2015). Selective chromatic contrast adaptation has also been shown as early as V1 (Engel and Furmanski, 2001), indicating examination of chromatic responses in early visual regions using fMRI would likely give insight into the extent and location of chromatic gain in anomalous trichromats. However, surprisingly little is known about the neural responses to color in color deficient observers. In particular there have not been studies of the

magnitude of neural responses in these observers using modern the neuroimaging techniques of fMRI.

In the present work, I examined both the mechanisms and sites of color compensation in anomalous trichromats using fMRI. These individuals provide an ideal "natural" experiment for understanding long-term plasticity and calibration in the human visual system, by comparing how cortical color responses are scaled between individuals who have a lifetime of exposure to similar visual worlds, but viewed through very different sensors.

Methods

Anomaloscope

The color vision of all participants was assessed using standard color screening techniques, including an anomaloscope. We used a Heidelbert-Multi-Color Anomaloskop (OCULUS, inc., Wetzlar, Germany). Participants were asked to match a pure orange light of 589 nm (+/- 2 nm, half-width of 10 nm) with a mixture of red (666 nm) and green (549 nm). The experimenter would select a mixture of red and green and ask the participants to attempt a match by only adjusting the brightness of the orange field. The experimenter would present different red-green mixtures until a participant's match range was determined. The match range was then used to calculate the anomaly quotient, which was used as the diagnostic criterion with the following rules: an AQ of less than 0.7 is protanomalous, an AQ range of 0.7 to 1.4 is color normal, and an AQ of greater than 1.4 is deuteranomalous, where 1 is an equal mixture of red and green, less than 1 is a higher proportion of red, and greater than 1 is a higher proportion of green. Dichromats are able

to match the orange test field to almost any red-green mixture by adjusting the brightness, and are thus easily differentiable from AT observers.

Participants

A total of 7 anomalous trichromats were recruited (3 deuteranomalous, 4 protanomalous, all male), ages 19-42. Color-normal control participants were all UNR students, ages 24-38. There were a total of 5 AT participants in experiment 1, and 5 controls (1 female). There were a total of 5 AT participants (3 repeated) in experiment 2 (ages 22-42), and 7 controls. One control participant was excluded due to erratic patterns of activation, leaving a total of 6 controls in experiment 2 (2 female, ages 24-38). We also ran 1 protanopic participant (male, age 32) in the threshold experiment and the high-attentional control task. Chromatic contrast detection thresholds were collected for all control and AT observers except 1 deuteranomalous participant.

Stimuli

Stimuli were presented on calibrated computer screens in testing rooms at UNR or the fMRI facility at Renown Health Hospital. Thresholds were collected using a SONY 20SE monitor, with displays controlled by a Cambridge Research Systems (Kent, UK) Visual Stimulus Generator (VSG) board, which allowed for high color resolution. Retinotopic and fMRI stimuli were displayed on a 32 in. SensaVue (85 Hz refresh rate) monitor (Invivo, Inc. Gainesville, FL) situated behind the scanner bore and viewed through a head-mounted mirror. The monitor's maximum visual field was 31 deg. by 19 deg. Both monitors were calibrated with a Photo Research (Syracuse, NY) PR655 spectroradiometer, with gun outputs linearized through lookup tables. Colors were adjusted to empirically equate their relative luminances for individual observers using the

minimum motion technique in order to reduce responses from potential luminance artifacts in the stimuli (Cavanagh et al., 1987). Participants were presented with counterphasing (1 Hz) radial sinewave gratings (0.28 c/deg., 14.5 deg field) defined by chromatic variations along the L/(L+M) or S/(L+M) cone opponent axes. The same stimuli were used in the scanner and during the threshold task (Mullen et al., 2007).

Each chromatic direction was shown at 4 levels of chromatic contrast for a total of 8 conditions (Figure 3). In experiment 1, the fixation was a black circle that randomly flickered at a jittered rate. In Experiment 2, the fixation was a series of numbers from 1 to 9 that were either black or white. The rate of change was also jittered. All stimuli were produced with the Psychophysics Toolbox (Brainard, 1997) for MATLAB (Mathworks Inc., Natick, MA).

Threshold Procedure

Participants were first adapted to a neutral gray background for 1 minute, and then completed a minimum motion task to determine isoluminance. The minimum motion stimuli consisted of alternating achromatic and chromatic radial sinusoidal gratings that were offset in phase. Participants were instructed to adjust the gratings until the stimuli no longer appeared to radiate inward or outward. Stimulus luminance was adjusted individually for each observer. We then used a 2-alternative forced choice task to measure contrast detection thresholds for detecting the gratings. The grating was shown in one of two temporal intervals signaled by beeps, and participants pressed a button to indicate at which interval the grating appeared. Contrast was varied in two randomly interleaved staircases (2 up, 1 down). Each staircase terminated after 10 reversals. The average of the last 8 reversals across each staircase was taken as the participant's

detection threshold. S/(L+M) and L/(L+M) contrast thresholds were collected during the same session, but across different runs.

fMRI Procedure

Scans were acquired on a Philips 3T Ingenia scanner using a 32-channel digital SENSE head coil (Philips Medical Systems, Best, Netherlands). Functional data for experiments 1 and 2 were obtained using T2*, echo Planar images (2 sec. TR, 180 volumes, voxel size 2.75x2.75x3 $mm^{3,}$ 36 slices, 55 ms inter slice time, 0 mm gaps). We used a block design where a run consisted of 16 blocks of 14 sec. each, interleaved with 8 sec. fixation-only gaps. Each stimulus type was presented twice per run, and the order was counterbalanced across runs. Each run took 360 sec. In experiment 1, all participants completed 8 runs. In experiment 2, participants completed 6 runs, reduced from experiment 1 to allow time for the collection of retinotopic mapping data. Anatomical scans occurred halfway through a session in order to reduce motion artifacts. Anatomical scans were collected using T1 weighted images (voxel size of 1x1x1 $mm^{3,}$ 30 sec transverse slices, 17 ms TE, 76^{o} flip angle, and 220 x 220 mm^{2} field of view). Anatomical and functional runs were collected in a single session for 12 participants, while the remaining 2 participants completed functional runs across two sessions.

Fixation Tasks

In Experiment 1, participants completed a simple fixation task to ensure that they were responsive and fixating during the stimulus presentation. The black fixation circle flickered at random, jittered intervals, and the participants were instructed to make a response when a change occurred. In order to further test and control for attentional

effects on the BOLD contrast response, in experiment 2 a second set of scans was collected using a more demanding fixation task (Kay and Yeatman, 2017). The fixation for this task was a number that randomly switched from black to white at a jittered rate, while also changing to a random number from 1 to 9. Participants were assigned a random pair of number conjunctions and asked to press a button when either number appeared (ex. black nine or white 3) (Goddard et al., 2010).

fMRI Data Preprocessing

All fMRI data were analyzed using Brainvoyager version 20.4 / BVQX 3.4 (Brain Innovation, Maastricht, The Netherlands). Automated intensity inhomogeneity corrections were performed on anatomical scans. ACPC landmarks were then identified by hand, and each anatomical scan was converted into Talairach coordinate space. Functional scans were slice-time corrected to the first slice, and high-pass temporal filtering (GLM-Fourier) was set at 2 cycles. 3D motion correction was aligned to the first volume of the functional run that occurred closest to the anatomical scan (the reference run), and runs with large movement artifacts were excluded from the group analysis. The reference run was used to generate transformation coordinates for functional to anatomical co-registration. We performed GLMs for each run, using fixation as the baseline (Goebel et al., 2006).

Retinotopic Mapping

In order to draw the retinotopic-based regions-of-interest, we reconstructed anatomical scans into inflated surfaces. Cortex surfaces were generated using BrainVoyager's advanced segmentation tool, which requires upsampling the anatomical data to 0.5x0.5.0.5 mm^3 resolution. This enables more accurate representations of white

matter/gray matter and gray matter/ CSF boundaries. Cortex meshes were then reconstructed based on identified boundaries. Following mesh reconstruction, we then downsampled back to the original resolution. Each mesh was inflated and smoothed. This process is outlined in BrainVoyager's QX User's Guide, Version 2.8 (Copyright 2014, Rainer Goebel).

Retinotopic mapping stimuli were flickering color and luminance checkerboard wedges (4 Hz), varying in polar angle, covering 45 deg. and spanning a total of 13.5 deg. of eccentricity. All participants completed at least two runs of retiniotopic mapping, one clockwise and one counter-clockwise (Swisher et al., 2007; Arcaro et al., 2009; Killebrew et al., 2015). Polar angle representations were obtained using BrainVoyager's Linear Correlation tool, using either 0 deg. or 180 deg. as reference and a cross correlation lag of 18. Two reference time course files were created for each polar angle run. The time courses were then projected onto an inflated, reconstructed surface. ROIs were defined using the reversals of activation, and drawn onto an inflated surface before being converted back into volume space. ROIs were identified using standard methods (Sereno et al., 1995; DeYoe et al., 1996; Engel et al., 1997; Press et al., 2001; Wade et al., 2002; Wandell et al., 2007; Arcaro et al., 2009). For reviews see (Wandell and Winawer, 2011) and (Silver and Kastner, 2009). An example of the drawn ROIs for one participant is shown in Figure 4.

Results

Thresholds

Thresholds were collected for 6 out of the 7 AT observers. Another AT observer's threshold data was excluded due to erratic staircase functions. This observer was also not included in experiment 2. We also excluded 1 control participant's data from both the thresholds and fMRI analysis, due to erratic BOLD responses. The contrast thresholds for 6 control and 5 AT observers are shown in Figure 5. As expected, the L/(L+M) thresholds were significantly higher for AT than control participants (AT mean = 9.39, AT SD = 1.81, Control mean = 1.29, Control SD = 0.25, $t(9) = 10.976$, $P<0.001$), while the S/(L+M) thresholds were not significantly different (AT mean = 1.77, AT SD = 0.35, Control mean = 1.71, Control SD = 0.57, $t(9) = 0.221$, $P = 0.830$).

Experiment 1: Simple Fixation

Experiment 1 was designed as an exploratory pilot study to assess the relative BOLD responses in color normal and anomalous observers. Scans were measured for a total of 10 participants (5 color normal, 2 deuteranomalous, and 3 protanomalous). For each participant, we performed a whole-brain GLM contrasting the responses to each chromatic grating relative to the interleaved gray field. Activation was limited to early visual areas, and the amount of activation was visually similar across groups. Two representative examples can be found in Figure 6. Retinotopic mapping was not performed for this initial set of data, so BOLD activation was sampled from an activation-based ROI in an early visual area. These results are shown in Figure 7, which compares the contrast response for normal and AT observers along the L-M and S axes. Symbols show the mean settings with standard errors for each group.

Fitting Contrast Response Function

In order to characterize how the responses varied with contrast, we fitted each data set with a standard contrast response function (Boynton et al., 1999).

$$R(c) = R_{max} * \frac{c^{(p+q)}}{\left(c^q + c_{50}^q\right)}$$

Where R is the measured BOLD response, c is the stimulus contrast, and p and q are exponents controlling the nonlinearity. These were fixed at p = 0.33 and q = 1.55 based on previous studies (Boynton et al,. 1999). We then varied Rmax (which controls the maximum response) and c50 (which controls the contrast at which the response reaches half the maximum) in order to find the values that provided a least-squares fit to the average responses for the control group. The fitted function is shown by the dashed gray line in Figure 8, and provided a good approximation to the observed BOLD amplitude.

To assess responses in the anomalous observers, the CRF function for the control participants was next rescaled by altering the effective contrast based on the threshold ratio for the anomalous trichromats (i.e. the function was re-evaluated with c reduced by the ratio of the normal to anomalous thresholds). If no compensation is occurring, the BOLD responses for the anomalous observers should follow the CRF adjusted for their threshold sensitivity (black dashed line, figure 8). Finally, we also fit the actual BOLD responses for the AT observers by varying the contrast scaling of the normals' CRF. Specifically, we again used the same parameters for the normal CRF fits, but varied the effective stimulus contrast to find the least squares fit to the anomalous responses. This estimate is shown by the red dashed line in Figure 8.

Statistical Comparisons

Based on this analysis, the effective L-M contrast for the AT observers was 44% of the controls for the BOLD responses, compared to 14% for the threshold ratios, a 3.1-fold difference. This implies that the observers were – on average – showing greater responses to the L-M stimuli than their thresholds predicted, yet weaker responses than the control group, or in other words, partial compensation for their cone sensitivity losses. This effect was evaluated by a two-way repeated measures ANOVA, where stimulus contrast was within subject, and group (control, AT, or threshold predicted) was between subjects. Error terms were based on the variability in the BOLD responses or thresholds. There were significant main effects of contrast ($F(1,12) = 48.56$, $P < 0.001$), as well as group ($F(2,12) = 13.234$, $P = 0.001$). There was no interaction effect ($F(2,12) = 1.915$, $P = 0.190$). Pairwise comparisons between groups revealed significant differences between AT vs. control (95% CI, $P = 0.025$), and AT vs. threshold predictions (CI 95%, $P = 0.024$) again suggesting that ATs showed partial but incomplete amplification of their L-M cone signals.

Experiment 2: High-Attentional Fixation

The next experiment was designed to evaluate how the BOLD responses in normal and AT observers varied with visual area, and also incorporated the more demanding fixation task in order to further minimize attentional influences on the BOLD response. Attention and task can strongly modulate the BOLD signal and alter the relationship between amplitude and contrast (Kay and Yeatman, 2017). The high-demand fixation task was therefore chosen to further isolate “bottom-up” processing in the color contrast responses. This experiment had a total of 12 participants (7 controls, 3

protanomalous, and 2 deuteranomalous). Five of the controls and 3 AT observers ran in both experiments, while the remaining participants were new. One control observer was excluded due to inconsistent BOLD activations across runs and possible motion artifacts, leaving a total of 11 participants (6 controls, 3 protanomalous, and 2 deuteranomalous). Runs with inconsistent or unusual activation were eliminated as well as runs where participants were either not responding or underperforming on the attention task. Using retinotopic mapping from independent scans, we performed a region-of-interest analysis for areas V1-V4. There was very little activation in areas V2d and V3d, so they are not included here. Figure 9 shows percent signal change averaged across participants. Error bars represent the standard error of the mean.

Comparisons across Experiments 1 and 2

Anomalous observers PA-01, DA-04, and PA-05 participated in both experiment 1 (low-attention fixation task), and experiment 2 (high-attention fixation task). In order to compare across these two tasks, functional runs were re-aligned with anatomical scans used for retinotopic mapping. Figure 10 shows the average BOLD responses for these 3 AT observers in both the high-attention fixation task and the low-attention fixation task for areas V1, V2v, V3v, and V4. We performed repeated measures ANOVAs for each of these regions, where contrast and attention tasks were within subject variables. Results in V1 do not show a significant effect of task ($F(1,2) = 0.068$, $P = 0.819$). The same was true in area V2v ($F(1,2) = 1.829$, $P = 0.309$), V3v ($F(1,2) = 0.778$, $P = 0.471$), and V4 ($F(1,2) = 2.492$, $P = 0.255$).

We repeated this analysis for control participants who ran in both studies. Following ROI analyses, 1 control subject's data did not show consistent contrast

responses and was thus excluded. Another control, was excluded from experiment 2 for similar reasons. Thus only 3 controls were included in this analysis (Figure 10). In area V1 there was not a significant effect of attention ($F(1,2) = 1.591$, $P = 0.334$). The same was true in V2v ($F(1,2) = 9.479$, $P = 0.091$), V3v ($F(1,2) = 6.052$, $P = 0.133$), and V4 ($F(1,2) = 0.342$, $P = 0.618$). These results indicate that top-down influences may not have differentially affect the BOLD responses of anomalous or control observers to chromatic stimuli in the two different fixation paradigms, though it should be noted that this conclusion is based on a small sample size.

Fitting CRF

We repeated the process from experiment 1 of fitting each data set with a CRF (Boynton et al., 1999), and scaling the controls' CRF by the threshold ratio to produce threshold predicted data. Figure 11 shows the results from the contrast response fit for the L-M stimuli. The scaling factor that best fit the AT data was 32% in area V1, 34% in area V2v, 25% in area V3v, and 32% in area V4. Each of these represents an approximate 2-fold change from the 14% scaling based on threshold predictions.

Statistical Comparisons within ROIs

ANOVA. We used the control data, the AT data, and the data derived from the threshold scaled CRF to run repeated measures ANOVAs, where one factor was group (control, AT, and threshold predicted data), and the other factor was contrast level (within subject). Like Experiment 1, this analysis again showed that BOLD responses were lower for ATs than controls in each visual area. However, unlike the preceding experiment, the differences between the observed and predicted AT responses in this case did not reach significance within any of the areas. The specific results were as follows.

The ANOVA for V1 resulted in a significant main effect for group ($F(2, 13) = 13.169$, $P = 0.001$) and contrast ($F(1,13) = 48.935$, $P < 0.001$), with no significant interaction ($F(2,13) = 1.103$, $P = 0.448$). The pairwise comparison between the AT group and the control group was significant (CI 95%, $P = 0.008$), but the comparison between the AT and the threshold predictions was not significant (CI 95%, $P = 0.09$).

In area V2v, there was a significant main effects of group ($F(2,13) = 8.909$, $P = 0.004$), and contrast ($F(1,13) = 26.652$, $P < 0.001$), and no significant interaction ($F(2,13) = 1.226$, $P = 0.325$). Control vs. AT was significant in area V2v (CI 95%, $P = 0.025$), but AT vs. threshold predictions was not significant (CI 95%, $P = 0.142$).

In area V3v, there was a significant effect of group ($F(2,13) = 6.899$, $P = 0.009$), and contrast ($F(1,13) = 37.891$, $P < 0.001$), with no interaction ($F(2,13) = 1.3$, $P = 0.306$). Control vs. AT was significant (CI 95%, $P = 0.015$), but AT vs. threshold was not significant (CI 95%, $P = 0.561$).

In area V4 there was a main effect of group ($F(2,13) = 9.068$, $P = 0.003$), and contrast ($F(20.554$, $P = 0.001$), with no interaction ($F(2,13) = 1.688$, $P = 0.233$). AT vs. control was significant (CI 95%, $P = 0.010$), but AT vs threshold predictions was not significant (CI 95%, $P = 0.326$).

Sign Test. As a second analysis, we also conducted a nonparametric sign test to determine whether the BOLD activations were larger than predicted by the thresholds for a significant proportion of times. For each region of interest, there were a total of 20 comparisons. However, this test was again not significant when assessed for each individual area [V1 ($P = 0.06$), V2v ($P = 0.41$), V3v ($P = 0.75$), or V4 ($P = 0.25$)].

ROIs as Within Subject Comparison

ANOVA. As an alternative analysis, we also included ROI as a within-subject variable and conducted a 3-way ANOVA. There was a significant effect of group ($F(2,13) = 15.230$, $P < 0.001$), ROI ($F(1,13) = 4.960$, $P = 0.044$), and contrast $F(1,13) = 41.883$, $P < 0.001$). There was a significant interaction between ROI and contrast ($F(1,13) = 16.983$, $P = 0.001$), but not between contrast and group ($F(2,13) = 1.146$, $P = 0.348$), or ROI and group ($F(2,13) = 2.520$, $P = 0.119$). In pairwise comparisons AT vs. control was significant (95% CI, $P = 0.003$), but threshold prediction vs. AT was not (95% CI, $P = 0.134$).

Sign Test. We also repeated the sign test, this time collapsing across regions. There were a total of 80 comparisons (AT observed BOLD – threshold predicted). 51 were positive, in this case suggesting that the observed BOLD responses were significantly larger than predicted ($P = 0.01$). Figure 13 shows responses from all subjects and ROIs plotted against the threshold predictions.

Comparisons of S/(L+M) Responses

Similar analyses were conducted to compare the responses of the control and AT groups for the S axis patterns. Figure 14 shows the mean BOLD responses for these S/(L+M) stimuli. We performed a repeated measures ANOVA comparing activation between the AT observers and the control participants, using ROI and contrast as within subject variables. There was not a significant effect of group ($F(1,9) = 1.055$, $P = 0.331$, 95% CI, $P = 0.331$) or ROI ($F(1,9) = 0.001$, $P = 0.982$). There was an effect of contrast ($F(1,9) = 32.55$, $P < 0.001$). Thus as expected AT observers were as sensitive as controls and showed similar BOLD activations for the S-cone varying stimuli.

Data Normalized to S/(L+M) Responses

A notable feature of both Experiments 1 and 2 was that the average BOLD responses of anomalous observers tended to be larger than in the control observers, though as noted above this difference was not statistically significant. Nevertheless, as a further comparison of the relative responses across observers, we normalized the L/(L+M) responses where 0 is the response at baseline and 1 is the response at S/(L+M) contrast 80, making L-M responses relative to S activation. Figure 15 shows this data, and the formula used for normalization is shown here:

$$normR = \frac{R_c - R_0}{R_{S80} - R_0}$$

As in previous analyses, we used the normalized data to fit a CRF to the mean control data, and derived a threshold prediction by scaling this CRF using the threshold ratio of controls to AT observers. The threshold ratio was 15% (also scaled to S/(L+M) responses). We then scaled the CRF of control observers to fit the AT data, deriving scaling factors for each region of interest. The scaling factor of control to AT was 24% in area V1, 35% in area V2v, 32% in area V3v, and 26% in area V4.

Statistical Comparisons within ROIs

ANOVA. We used this data to run two-way ANOVAs for each ROI, again using the AT data, the control data, and data derived from the CRF scaled by the AT threshold ratio. Overall, normalizing reduced effect sizes and variability, and again did not reveal significant group differences between the observed BOLD responses for ATs and the threshold-predicted responses.

The ANOVA in V1 showed a significant main effect of group ($F(2,13) = 20.679$, $P < 0.001$) and contrast ($F(1,13) = 39.742$, $P < 0.001$). There was not a significant interaction ($F(2,13) = 1.665$, $P = 0.227$). There was a significant difference between control vs. AT (95% CI, sig. = 0.001), but no difference between AT and the threshold predictions (95% CI, sig. = 0.248).

In area V2v, there was a main effect of group ($F(2,13) = 7.798$, $P = 0.006$), and contrast ($F(2,13) = 23.190$, $P < 0.001$), with no interaction ($F(2,13) = 1.189$, $P = 0.335$). AT vs. control was significantly different (95% CI, $P = 0.020$), and AT vs. threshold predictions was not (95% CI, $P = 0.282$).

In area V3v, there was a main effect of group ($F(2,13) = 4.072$, $P = 0.042$), and contrast ($F(1,13 = 17.364$, $P = 0.001$), with no interaction ($F(2,13) = 0.656$, $P = 0.535$). Control vs. AT was not significant (95% CI, $P = 0.067$), and AT vs. threshold predictions was not (95% CI, $P = 0.503$).

In area V4, there was not a main effect of group ($F(2,13) = 3.387$, $P = 0.065$), but there was an effect of contrast ($F(1,13) = 17.176$, $P = 0.001$), and no interaction ($F(2,13) = 1.279$, $P = 0.311$).

Sign Test. We again also ran a sign test (AT data – threshold predictions) in each region of interest. There was a significant difference in area V1, ($P = 0.01$), but not in area V2v ($P = 0.13$), area V3v ($P = 0.41$), or area V4 ($P = 0.41$).

ROIs as Within Subject Comparison

ANOVA. We repeated the ANOVA by adding ROI as a within subject comparison. There was a significant effect of group ($F(2,13) = 7.787$, $P = 0.007$), ROI ($F(1,13) = 26.033$, $P < 0.001$), and contrast ($F(1,13) = 27.658$, $P < 0.001$), with a

significant interaction between ROI and contrast (F(1,13) = 7.558, P = 0.017), but not ROI and group (F(2,13) = 2.075, P = 0.165). In Pairwise comparisons there was a significant difference between AT observers and controls (CI 95%, P = 0.015), but not between AT observers and threshold predictions (CI 95%, P = 0.427).

Sign Test. In a sign test for the normalized data collapsed across ROIs, a total of 40 out of 80 comparisons were positive, P = 0.54.

Comparisons of Scaling Factors.

Paired t-test. As a still further assessment of compensation, we used data derived from scaling the CRF fit to the mean L/(L+M) activation of control participants to fit the AT observers. This scaling represents the magnitude of difference between the control and AT observers. We performed a paired-samples t-test in each ROI comparing the scaling derived from the BOLD signal, and the ratio of L-M threshold for controls vs. AT (mean = 0.14, SD = 0.02). There was not a significant difference between the AT scaling and the threshold ratio in V1 (mean = 0.40, SD = 0.30, t(4) = 1.844, P = 0.139), V2v (mean = 0.46, SD = 0.38, t(4) = 1.915, P = 0.128), V3v (mean = 0.41, SD = 0.37, t(4) = 1.558, P = 0.194), or V4 (mean = 0.46, SD = 0.53, t(4) = 1.345, P = 0.250). We repeated this test, averaging across regions-of-interest (mean = 0.43, SD = 0.31, t(4) = -2.032, P = 0.112).

We repeated the paired t-test using the data normalized relative to S/(L+M) responses. Again, there was not a significant difference between the scaling of BOLD signal from control to AT observers, and the threshold ratios of controls to AT (mean = 0.15, SD = 0.033) in area V1 (mean = 0.29, SD = 0.24, t(4) = 1.554, P = 0.195), V2v (mean = .58, SD = 0.61, t(4) = 1.080, P = 0.341), V3v (mean = .80, SD = 1.33, t(4) =

1.080, P = 0.341), V4 (mean = .64, SD = 0.90, t(4) = 1.167, P = 0.306), or when averaged across ROIs (mean = 0.58, SD = 0.75, t(4) = -1.241, P = 0.283).

We also derived ratios of S/(L+M) vs. L/(L+M) for the all participants' thresholds and the BOLD activation in each ROI. The BOLD ratios were derived from scaling the S/(L+M) CRF to fit the L/(L+M) data, providing a scaling factor of the relative activation of S to L-M. Figure 17 shows this data in log units. We performed t-tests on these ratios. There was a significant difference between the threshold ratios of control participants (mean = 1.35, SD = 0.46) and anomalous trichromats (mean = 0.19, SD = 0.04, t(9) = -5.59, P < 0.001). The same t-test was performed for areas V1 (AT mean = 0.25, SD = 0.07, control mean = 1.09, SD = 0.58, t(9) = -3.149, P = 0.012), V2v (AT mean = 0.54, SD = 0.29, control mean = 0.98, SD = 0.58, t(9) = -1.906, P = 0.089), V3v (AT mean = 0.422, SD = 0.38, control mean = 0.82, SD = 0.49, t(9) = -1.466, P = 0.177), and V4 (AT mean = 0.45, SD = 0.43, control mean = 1.09, SD = 1.07, t(9) = -1.237, P = 0.247). There was a significant difference between the two groups at threshold and in area V1, but there was no significant difference in areas V2v, V3v, and V4. This indicates a larger difference between AT and control participants at threshold and in area V1, with groups becoming more similar in V2v, V3v, and V4.

Basis for changes in the CRF

In the preceding analyses we estimated the AT observer's CRF by simply rescaling the effective contrast of the stimulus for the CRF fit to the mean of the color-normal observers. However, the CRF could vary in different ways, and in particular could reflect a change in specific parameters of the function. For example, variations in R_{max} represent a vertical rescaling of the response vs. contrast function, while variations in C_{50}

reflect a horizontal rescaling. To assess which shift better described the relationship between the control and AT responses, we refit the AT responses but this time only varied the value of R_{max} or C_{50} in the control CRF to determine which provided a better account of the individual AT responses. This was done for each individual and each visual area, as well as for the two different normalization criteria (i.e. normalizing the AT response relative to the control for L-M and S, or normalizing L-M relative to S). Across all conditions (5 observers x 4 visual areas times 2 normalized data sets), the average rms error in the fitted AT responses based on rescaling contrast (as we did previously) was 0.147. This was nearly identical to the fits obtained by varying R_{max} (0.146) since this has very similar effects on the CRF. Alternatively, the rms error was reduced to 0.125 when we instead varied only C_{50}. To assess whether this difference was significant, we again used a sign test to evaluate how many of the 40 comparisons resulted in a better fit for C_{50} than R_{max}. Excluding ties, this resulted in 26/32 cases where the fit was better for the parameter, a proportion that was highly significant (P = 0.0005). The SS error of the scaling fit vs. the rMax or c50 fits can be seen in Figure 18. These results are important in determining potential mechanisms that could underlie compensatory changes in AT individuals, with vertical rescaling associated with responses gain mechanisms, and horizonatal shifts associated with contrast gain.

Individual Observers

The comparisons thus far were based on mean differences between groups, and have low power because of the small sample sizes. In the next set of analyses we instead examined the results for individual observers to assess whether some might exhibit evidence for compensation.

Sign Tests. We first performed sign tests for each individual, this time collapsing across ROIs due to the visually consistent patterns across regions in individual participants. For each participant there were a total of 16 comparisons (4 contrasts x 4 ROIs). We did this first for the responses to L/(L+M) contrast (Figure 12). For 3 subjects were there was not a significant difference between their BOLD responses and their threshold predicted data: PA-01 ($P = 0.96$), DA-02 ($P = 0.77$), and DA-04 ($P = 0.11$). However, for two subjects there was a significant effect: PA-03 ($P < 0.01$) and PA-05 ($P < 0.01$).

We repeated this test using each individual's S-normalized responses (Figure 16). Again there were 3 subjects with no significant effect: PA-01 ($P = 0.99$), DA-02 (0.99), and PA-05 (0.99). There were also 2 subjects with significant effects: PA-03 ($P < 0.01$) and DA-04 ($P = 0.01$).

Notably, 1 subject who showed a significant effect of activation to L-M stimuli, no longer showed compensation after the data was normalized to his S-axis response (subject PA-05). Another subject did not show an effect of L-M activation, but had an effect following data normalization (subject DA-04).

Dichromatic Observer

As a control condition, we also tested one dichromatic (protanopic) participant in the fMRI paradigm of Experiment 2. PN-01's data can be seen in Figure 19, where it is compared to the average responses for the anomalous trichromats. PN-01's thresholds were 9.14 for the L/(L+M) stimuli and 2.02 for the S/(L+M). These are similar to the thresholds we obtained for some of the anomalous observers, though he was diagnosed as a protanope based on his anomaloscope settings, in which he accepted all red-green ratios

as matches. As expected, the M-cone based BOLD activation for this dichromatic observer was substantially weaker than for the ATs, yet surprisingly, his S responses were also lower. There were no regions where this participant showed a monotonic contrast response function for the L-M stimuli, and the magnitude of the response at the highest contrast (80) was no greater than the response to the lowest contrast (10). Thus we were unable to fit a reasonable CRF to his data. Overall then, PN-01's data showed relatively little activation to the L/(L+M) stimuli. This provides evidence that our stimuli were specifically targeting the chromatic responses in the normal and anomalous observers, and thus were not driven simply by potential luminance artifacts in the gratings.

Discussion

To summarize, we examined evidence for cortical amplification of chromatic signals in anomalous trichromats in two different experiments comparing fMRI BOLD signals in normal and anomalous observers in early visual areas. The two experiments differed primarily in the attentional task and in the localization of the responses to different retinotopically defined areas. The first preliminary experiment indicated robust compensation for the L-M chromatic response, with the BOLD responses on average 3 times higher than predicted by the anomalous subjects' threshold sensitivity. The second experiment revealed weaker effects, with an average gain of 2 times the responses predicted by the thresholds. The weaker effects for the second condition provide more mixed evidence for the compensation hypothesis – with some of the analyses suggesting significant amplification while most failing to reach significance. Both experiments also clearly agreed in showing that the anomalous responses to the L-M stimuli were

significantly weaker than in the normal trichromats. If compensation does occur, these seem to strongly indicate that it is not complete, at least for the specific conditions and tasks we examined. Nevertheless the two experiments together are suggestive that the chromatic signals for the anomalous observers do show partial compensation for the weakened chromatic signals provided by their cones.

It should be emphasized that there are two limitations to this analysis of the group differences. First, we were constrained to a very small sample size of roughly 5 observers per group, and this obviously limited the power of the group comparisons. Second, there were very large differences between the observers, suggesting simple group comparisons may not be appropriate. In particular it appeared that only some of the individuals showed significant compensation, and this was true for the two different assumptions we used for scaling their responses (relative to the controls or to their own S-axis contrast response). Analyses of the individual data suggest that out of our 5 AT participants, 2 showed strong compensation, 2 showed no compensation, and 1 showed compensation in only some regions. These results were not significantly augmented by the addition of a high-attentional control, indicating that these effects are not due to top-down influences. Our results suggest that some but not all AT observers are compensating for their reduced sensitivity to chromatic variations through a gain mechanism that likely exists as early as V1.

It is well recognized that anomalous trichromats are a heterogeneous population. For example, some ATs show biased anomaloscope matches yet very narrow matching ranges. These differences are also evident in the previous studies testing for compensation. In Regan & Mollon's study (Regan and Mollon, 1997), out of 17 AT

participants, 2 showed patterns that were nearly identical to the color normal observers. However, unlike our data, the two that showed a great deal of compensation were deuteranomalous, and performed better on color vision tasks than most anomalous trichromats. The participants who showed compensation in our task were not predicted by their chromatic thresholds, and included a deuteranomalous and a protanomolous observer. This is consistent with results from Boehm et al. (Boehm et al., 2014), who showed that some participants with very poor sensitivity still showed evidence of compensation in a subjective measure of color appearance. Boehm et al. also showed that, though on average AT participants' were much closer to color normals than their thresholds would predict, however 6 out of 14 AT observers fell below 2 standard deviations from the normal observers.

The reasons for the differences between observers are not understood. It is possible that chromatic compensation only occurs in some anomalous trichromats, yet a further possibility is that some anomalous trichromats would show evidence of compensation to a more specialized set of stimuli. In a multidimensional scaling procedure, Bosten et al. used specialized stimuli that were differentiable to observers with red-green color deficits, but were metameric to color normal participants. The color spaces derived from the scaling values were truly two dimensional for AT observers, indicating that anomalous trichromats might be expanding color differences via neural gain, but this expansion isn't necessarily evident using stimuli selected for color normal observers (Bosten et al., 2005).

Many of the observers from the present study also participated in a behavioral suprathreshold contrast matching task as part of a separate study (Vanston et al., 2017).

In this task, observers were presented with two consecutive gratings, and responded in a forced choice task to indicate which had a higher contrast, using an achromatic grating as a reference and either an L-M or S grating as the test. Surprisingly, in this task the suprathreshold matches were on average predicted by chromatic thresholds. These results can be seen in Figure 20. In addition, subjects that showed compensation in their BOLD signals did not have correspondingly stronger L-M percepts in their suprathreshold contrast matches. A further result of Vanston et al. is that the AT observers required higher contrast of both the S and the L-M gratings in order to match the reference achromatic grating. This differs from our findings that AT observers had higher activation to the S-axis stimuli, indicating increased sensitivity compared to color normal observers. However, our study did not present participants with achromatic gratings. Based on previous work, it would be expected that responses to achromatic gratings would not differ significantly from chromatic gratings (Mullen et al., 2007; Mullen et al., 2015), but this has not been examined in anomalous trichromats. In any case, the basis for these differences between the relative strength of chromatic signals we measured behaviorally and neurally remain uncertain.

Our ROI results are consistent with previous work suggesting that there are strong chromatic responses in early visual regions (Johnson et al., 2001, 2004; Johnson and Mullen, 2016). As in Mullen et al. 2007, (Mullen et al., 2007), chromatic responses for L/(L+M) were similar to responses to S/(L+M) stimuli in areas V1-hV4. Additionally, in participants that showed strong compensation, there was compensation as early as V1, leaving open the possibility that compensation is occurring at some earlier, pre-cortical stage, including the retina or LGN. However, when comparing relative activation

between S responses and L-M responses, there was a significant difference across groups at area V1, and no significant difference in areas V2v, V3v, and V4. This might suggest that while individual observers who showed compensation were showing it as early as V1, the overall difference in relative activation across AT and control observers was more similar in later visual areas, consistent with later sites. Again, though, our analyses were mixed in that some group differences pointed to compensation while others did not. Lutze et al. has previously shown that achromatic contrast gain in trichromats and dichromats is equivalent in parvocellular pathway through LGN (Lutze et al., 2006). From this they argued that the geniculate gain was genetically determined and thus not affected by experience. However, their results could also be explained by assuming that trichromats and dichromats are exposed to the same range of achromatic signals, and thus under similar states of cortical adaptation to luminance contrast. This would allow the possibility that observers with different chromatic sensitivities could in fact adapt to adjust for these differences.

Other work has also pointed to potentially later sites of compensation. In a very recent study, Rabin et al. (Rabin et al., 2017) observed patterns consistent with compensation in anomalous trichromats by measuring visual evoked potentials to cone-isolating stimuli. They found that VEP amplitudes were very similar to VEPs of color normal participants, but only when the stimuli were viewed binocularly. Again the reason why compensation might require binocular mechanisms is obscure, but this could indicate that a more central mechanism with greater interocular transfer or weighting relative responses from different locations in the two eyes' visual fields is necessary to see chromatic compensation. However, these results were not measured at multiple

contrast levels, which could indicate that his measure of compensation was indicative of a saturation of chromatic signal. In this regard we note again that our measurements were made with binocular viewing.

The results of these studies are relevant not only to color but to understanding adaptation and neural calibration more generally, for the patterns and mechanisms of plasticity appear to be very similar across sensory levels and modalities. Testing these effects in color deficiencies are again an ideal context because these deficiencies are prevalent and provide a natural experiment for gauging how far and in what ways the visual system can overcome a sensitivity loss. The results will also have important clinical implications. Recent studies have demonstrated the potential of gene therapy for "curing" color blindness, by introducing a novel photopigment gene into dichromatic primates (Mancuso et al., 2009). These primates expressed the photopigment in their receptors and became trichromatic as assessed by color screening tests. Clinical trials of this gene therapy are planned to begin soon, but virtually nothing is known about how or how well individuals will adapt to the changes in their color vision. Future studies in this area will provide novel first steps in understanding and predicting the dynamics of adjustments for a change in color sensitivity.

References

Arcaro MJ, McMains SA, Singer BD, Kastner S (2009) Retinotopic organization of human ventral visual cortex. The Journal of Neuroscience 29:10638-10652.

Asenjo AB, Rim J, Oprian DD (1994) Molecular determinants of human red/green color discrimination. Neuron 12:1131-1138.

Barbur J, Rodriguez-Carmona M, Harlow J, Mancuso K, Neitz J, Neitz M (2008) A study of unusual Rayleigh matches in deutan deficiency. Visual neuroscience 25:507-516.

Boehm AE, MacLeod DIA, Bosten JM (2014) Compensation for red-green contrast loss in anomalous trichromats. Journal of vision 14.

Bosten J, Robinson J, Jordan G, Mollon J (2005) Multidimensional scaling reveals a color dimension unique to 'color-deficient'observers. Current Biology 15:R950-R952.

Boynton GM, Demb JB, Glover GH, Heeger DJ (1999) Neuronal basis of contrast discrimination. Vision research 39:257-269.

Brainard DH (1997) The psychophysics toolbox. Spatial vision 10:433-436.

Cavanagh P, MacLeod DI, Anstis SM (1987) Equiluminance: spatial and temporal factors and the contribution of blue-sensitive cones. JOSA A 4:1428-1438.

Chaparro A, Stromeyer 3rd C, Huang E, Kronauer R, Eskew Jr R (1993) Colour is what the eye sees best. Nature 361:348.

Crognale M, Teller D, Motulsky A, Deeb S (1998) Severity of color vision defects: electroretinographic (ERG), molecular and behavioral studies. Vision Research 38:3377-3385.

Crognale MA, Switkes E, Rabin J, Schneck ME, Hægerström-Portnoy G, Adams AJ (1993) Application of the spatiochromatic visual evoked potential to detection of congenital and acquired color-vision deficiencies. JOSA A 10:1818-1825.

Deeb S, Lindsey D, Hibiya Y, Sanocki E, Winderickx J, Teller D, Motulsky A (1992) Genotype-phenotype relationships in human red/green color-vision defects: molecular and psychophysical studies. American journal of human genetics 51:687.

Delahunt PB, Webster MA, Ma L, Werner JS (2004) Long-term renormalization of chromatic mechanisms following cataract surgery. Visual neuroscience 21:301-307.

DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J (1996) Mapping striate and extrastriate visual areas in human cerebral cortex. Proceedings of the National Academy of Sciences 93:2382-2386.

Engel S, Zhang X, Wandell B (1997) Colour tuning in human visual cortex measured with functional magnetic resonance imaging. Nature 388:68-71.

Engel SA, Furmanski CS (2001) Selective adaptation to color contrast in human primary visual cortex. Journal of Neuroscience 21:3949-3954.

Goddard E, Mannion DJ, McDonald JS, Solomon SG, Clifford CWG (2010) Combination of subcortical color channels in human visual cortex. Journal of vision 10:25.

Goebel R, Esposito F, Formisano E (2006) Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Human brain mapping 27:392-401.

Johnson EN, Mullen KT (2016) Color in the Cortex. In: Human Color Vision, pp 189-217: Springer.

Johnson EN, Hawken MJ, Shapley R (2001) The spatial transformation of color in the primary visual cortex of the macaque monkey. Nature neuroscience 4:409-416.

Johnson EN, Hawken MJ, Shapley R (2004) Cone inputs in macaque primary visual cortex. Journal of Neurophysiology 91:2501-2514.

Kay KN, Yeatman JD (2017) Bottom-up and top-down computations in word-and face-selective cortex. eLife 6:e22341.

Killebrew K, Mruczek R, Berryhill ME (2015) Intraparietal regions play a material general role in working memory: Evidence supporting an internal attentional role. Neuropsychologia 73:12-24.

Kuriki I, Sun P, Ueno K, Tanaka K, Cheng K (2015) Hue selectivity in human visual cortex revealed by functional magnetic resonance imaging. Cerebral Cortex:bhv198.

Lutze M, Pokorny J, Smith VC (2006) Achromatic parvocellular contrast gain in normal and color defective observers: Implications for the evolution of color vision. Visual neuroscience 23:611-616.

MacLeod D (2002) Color discrimination, color constancy and natural scene statistics. Journal of Vision 2:15-15.

MacLeod DI, Hayhoe M (1974) Three pigments in normal and anomalous color vision. JOSA 64:92-96.

Mancuso K, Hauswirth WW, Li Q, Connor TB, Kuchenbecker JA, Mauck MC, Neitz J, Neitz M (2009) Gene therapy for red–green colour blindness in adult primates. Nature 461:784-787.

McDermott KC, Webster MA (2012) The perceptual balance of color. Josa a 29:A108-A117.

Merbs SL, Nathans J (1992a) Absorption spectra of the hybrid pigments responsible for anomalous color vision. Science 258:464-467.

Merbs SL, Nathans J (1992b) Absorption spectra of human cone pigments. Nature 356:433.

Mullen KT, Chang DH, Hess RF (2015) The selectivity of responses to red-green colour and achromatic contrast in the human visual cortex: an fMRI adaptation study. European Journal of Neuroscience 42:2923-2933.

Mullen KT, Dumoulin SO, McMahon KL, De Zubicaray GI, Hess RF (2007) Selectivity of human retinotopic visual cortex to S-cone-opponent, L/M-cone-opponent and achromatic stimulation. European Journal of Neuroscience 25:491-502.

Müller M, Cavonius C, Mollon J (1991) Constructing the color space of the deuteranomalous observer. In: Colour vision deficiencies X, pp 377-387: Springer.

Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS (1986) Molecular genetics of inherited variation in human color vision. Science 232:203-210.

Neitz J, Neitz M (2011) The genetics of normal and defective color vision. Vision research 51:633-651.

Neitz J, Neitz M, Kainz PM (1996) Visual Pigment Gene Structure and the Severity of Color Vision Defects. In, pp 801-804. Science.

Neitz M, Neitz J (2000) Molecular genetics of color vision and color vision defects. Archives of Ophthalmology 118:691-700.

Press WA, Brewer AA, Dougherty RF, Wade AR, Wandell BA (2001) Visual areas and spatial summation in human visual cortex. Vision research 41:1321-1332.

Rabin J, Kryder A, Lam D (2017) Binocular facilitation of cone-specific visual evoked potentials in colour deficiency. Clinical and Experimental Optometry.

Rabin JC, Kryder AC, Lam D (2016) Diagnosis of Normal and Abnormal Color Vision with Cone-Specific VEPs. Translational vision science & technology 5:8-8.

Regan BC, Mollon JD (1997) The relative salience of the cardinal axes of colour space in normal and anomalous trichromats. In: Colour vision deficiencies XIII, pp 261-270: Springer.

Rieke F, Rudd ME (2009) The challenges natural images pose for visual adaptation. Neuron 64:605-616.

Sanocki E, Teller DY, Deeb SS (1997) Rayleigh match ranges of red/green color-deficient observers: psychophysical and molecular studies. Vision research 37:1897-1907.

Sereno M, Dale A, Reppas J, Kwong K (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268:889.

Shevell S, He J, Kainz P, Neitz J, Neitz M (1998) Relating color discrimination to photopigment genes in deutan observers. Vision research 38:3371-3376.

Silver MA, Kastner S (2009) Topographic maps in human frontal and parietal cortex. Trends in cognitive sciences 13:488-495.

Stringham NT, Sabatinelli D, Stringham JM (2013) A potential mechanism for compensation in the blue—yellow visual channel. Frontiers in human neuroscience 7.

Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. The Journal of neuroscience 27:5326-5337.

Vanston JE, Tregillus KEM, Crognale MA (2017) Perceptual compensation in anomalous trichromats? In. Journal of Vision.

von der Twer T, MacLeod DI (2001) Optimal nonlinear codes for the perception of natural colours. Network: Computation in Neural Systems 12:395-407.

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

Wade AR, Brewer AA, Rieger JW, Wandell BA (2002) Functional measurements of human ventral occipital cortex: retinotopy and colour. Philosophical Transactions of the Royal Society of London B: Biological Sciences 357:963-973.

Wandell BA, Winawer J (2011) Imaging retinotopic maps in the human brain. Vision research 51:718-737.

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366-383.

Webster MA (2011) Adaptation and visual coding. Journal of Vision 11.

Webster MA, Leonard D (2008) Adaptation and perceptual norms in color vision. Josa a 25:2817-2825.

Webster MA, Juricevic I, McDermott KC (2010) Simulations of adaptation and color appearance in observers with varying spectral sensitivity. Ophthalmic and Physiological Optics 30:602-610.

Werner JS, Schefrin BE (1993) Loci of achromatic points throughout the life span. Josa a 10:1509-1516.

Figures

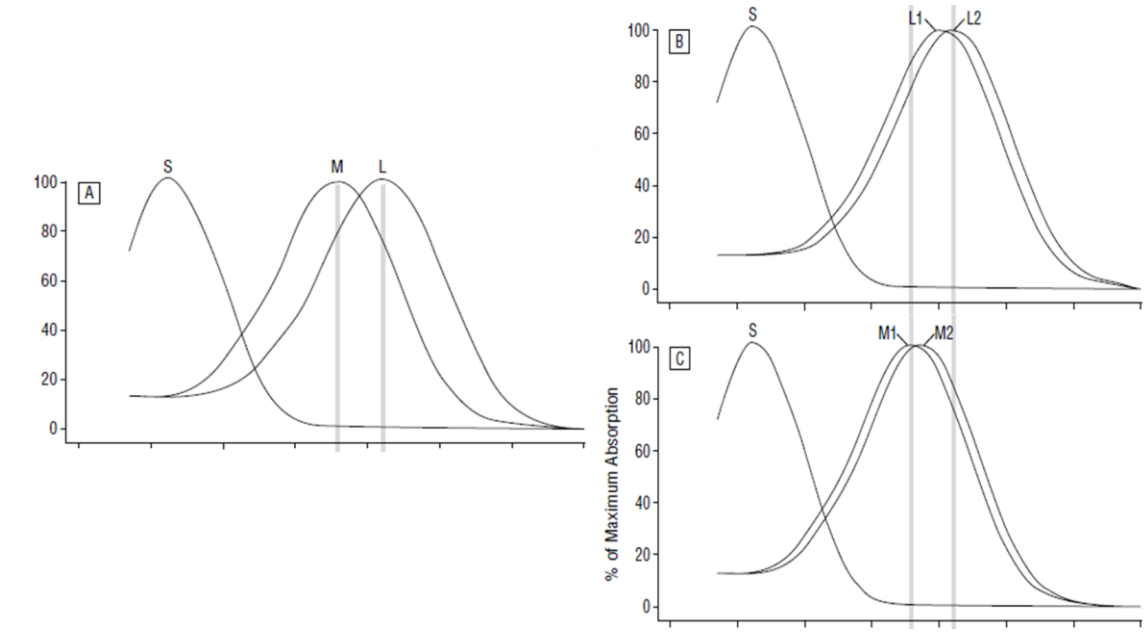

Figure 1. *Example of shifted cones in anomalous trichromacy, from (M. Neitz & Neitz, 2000). A) The sensitivity spectra for each of the three cone types in a trichromat. B & C) The three cone types present in deuteranomalous and protanomalous individuals.*

Figure 2. *A) An original color image. B) The image as seen by a protanomalous observer assuming adaptation only in the cones to match the average spectrum. C) The image assuming contrasts are rescaled by post-receptoral gain. Figure from (Webster et al., 2010).*

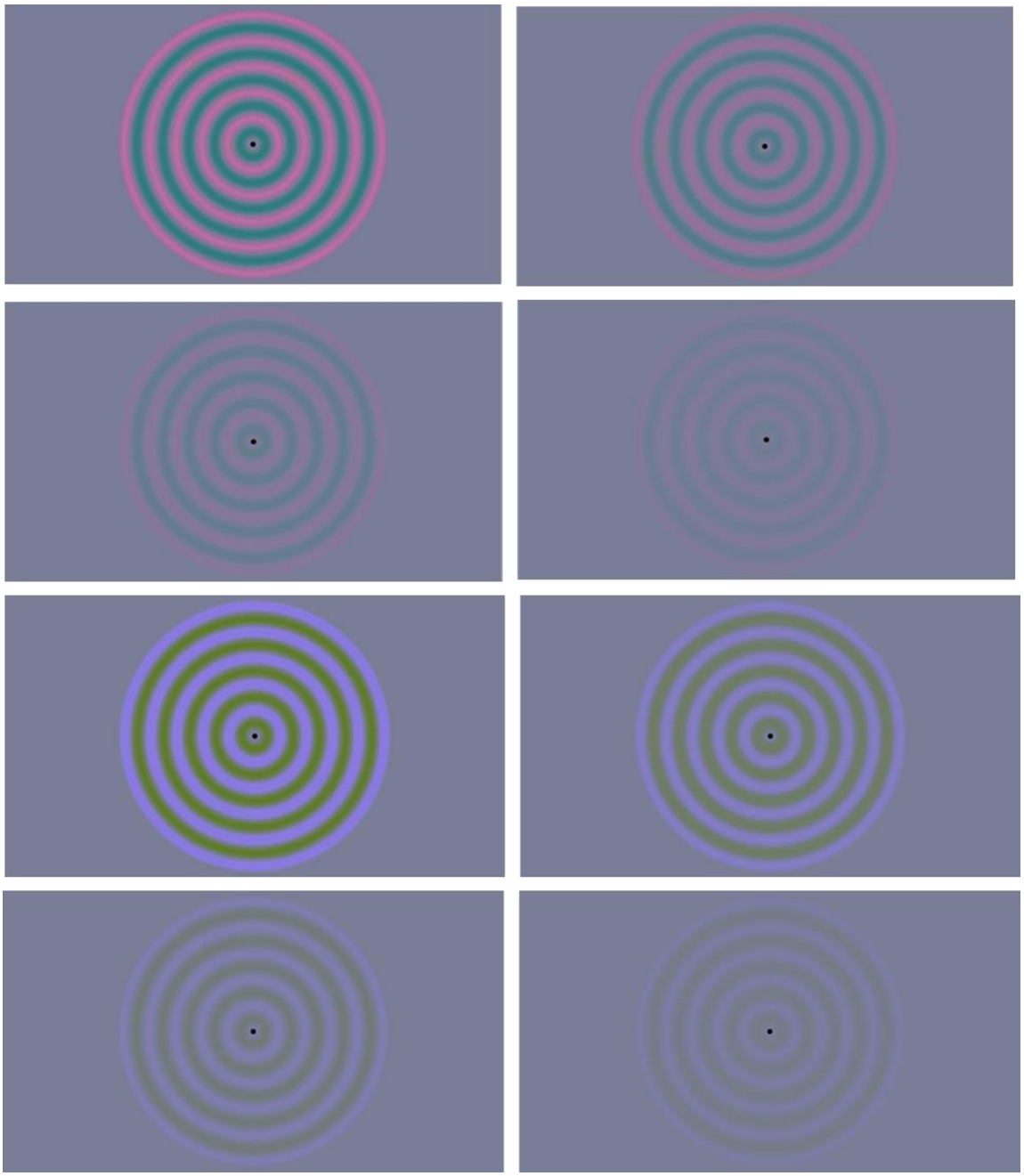

Figure 3. *fMRI stimuli, top two rows are the L/L(+M) gratings and the bottom two are the S/(L+M) gratings. Each was presented at 4 contrast levels, 80, 40, 20, and 10 in units of nominal multiples of threshold. These radial sine wave gratings were also used in collecting detection thresholds.*

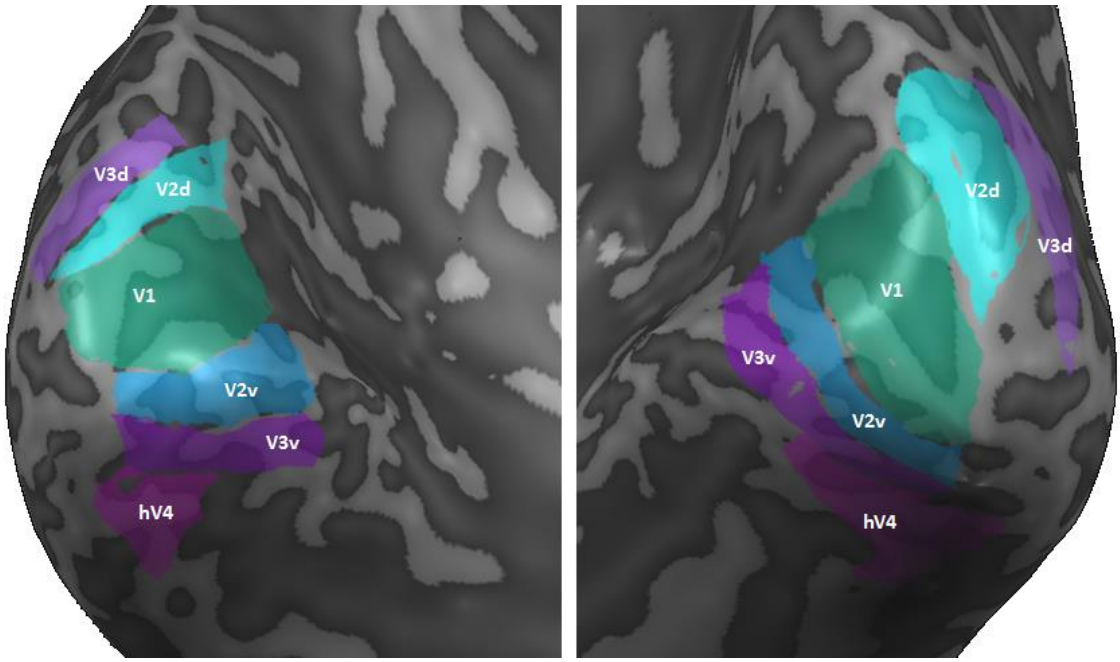

Figure 4. *An example of ROIs for 1 observer (C-02) on an inflated brain.*

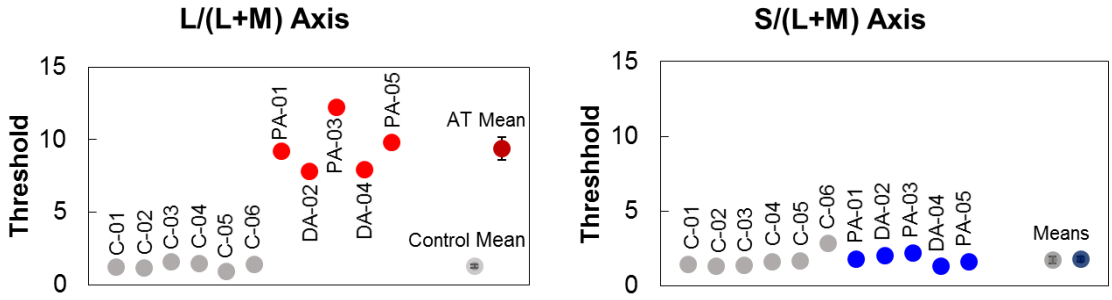

Figure 5. *Raw threshold results for the two chromatic axes, represented in nominal multiples of threshold. The x-axis is subject number, with control participants in gray.*

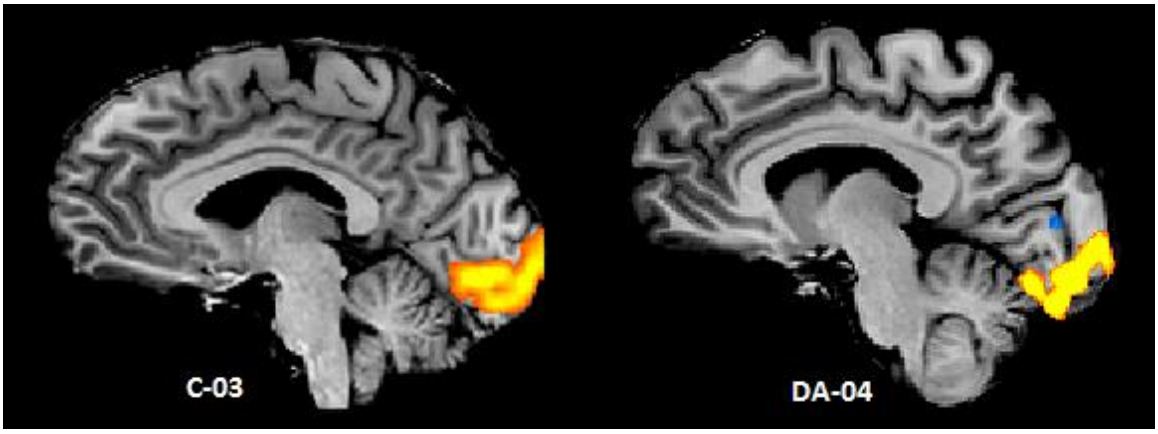

Figure 6. *Whole brain GLM (contrast: all conditions greater than fixation), for two representative observers. C-03, control observer. DA-04, deuteranomalous observer.*

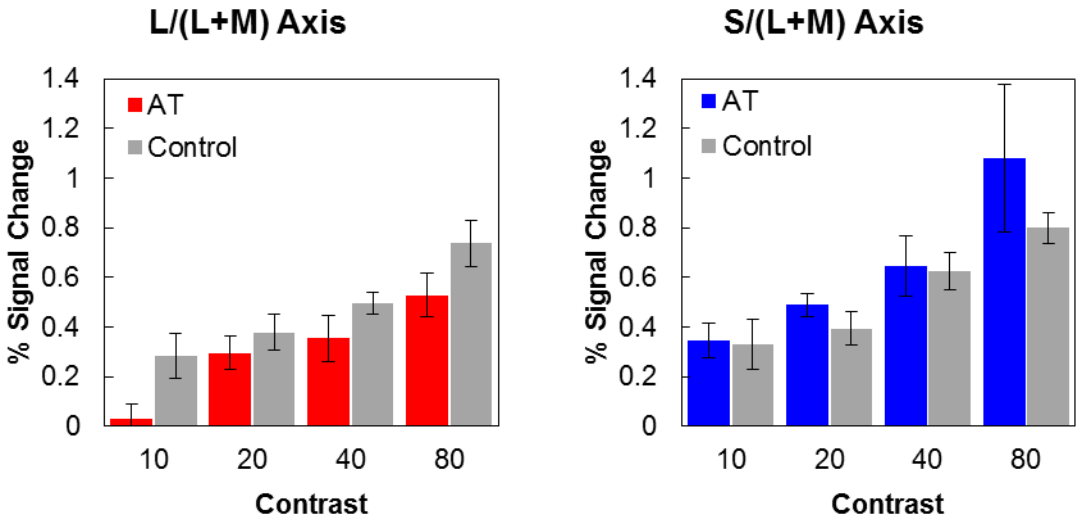

Figure 7. *fMRI results from Exp. 1, Simple Fixation task. Shown here is all activity greater than baseline using a GLM from an activation-based ROI in early visual cortex, and averaged across observers. Error bars represent the standard error of the mean.*

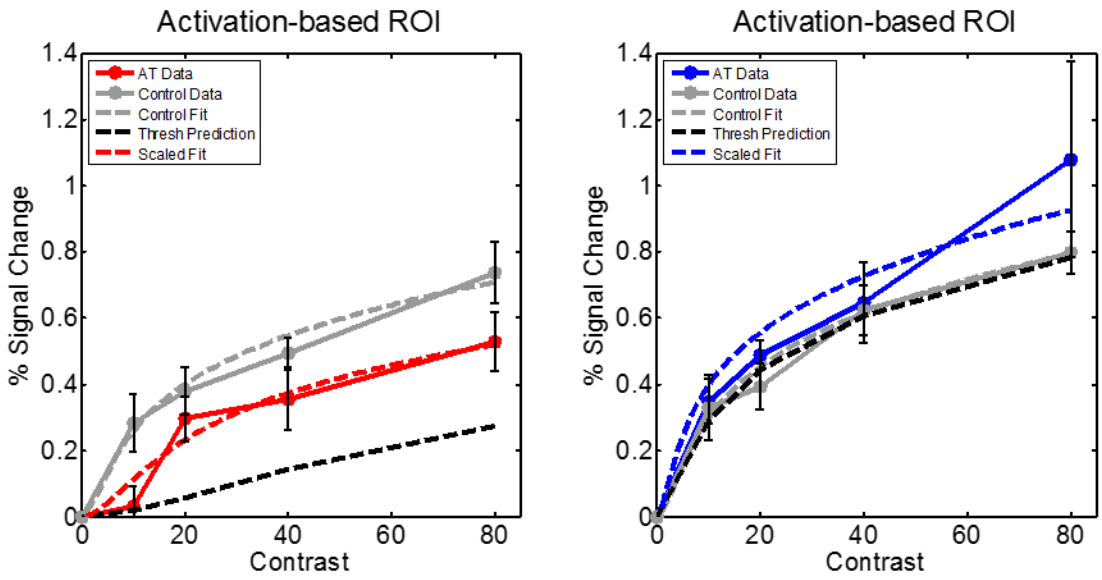

Figure 8. *Solid lines are the BOLD responses from Exp. 1, simple fixation task. The dashed lines are CRFs fitted to the data. The black, dashed line is the CRF of the control data, scaled to the threshold ratio of control to AT. Red are the L/(L+M) data, and blue are the S/(L+M) data.*

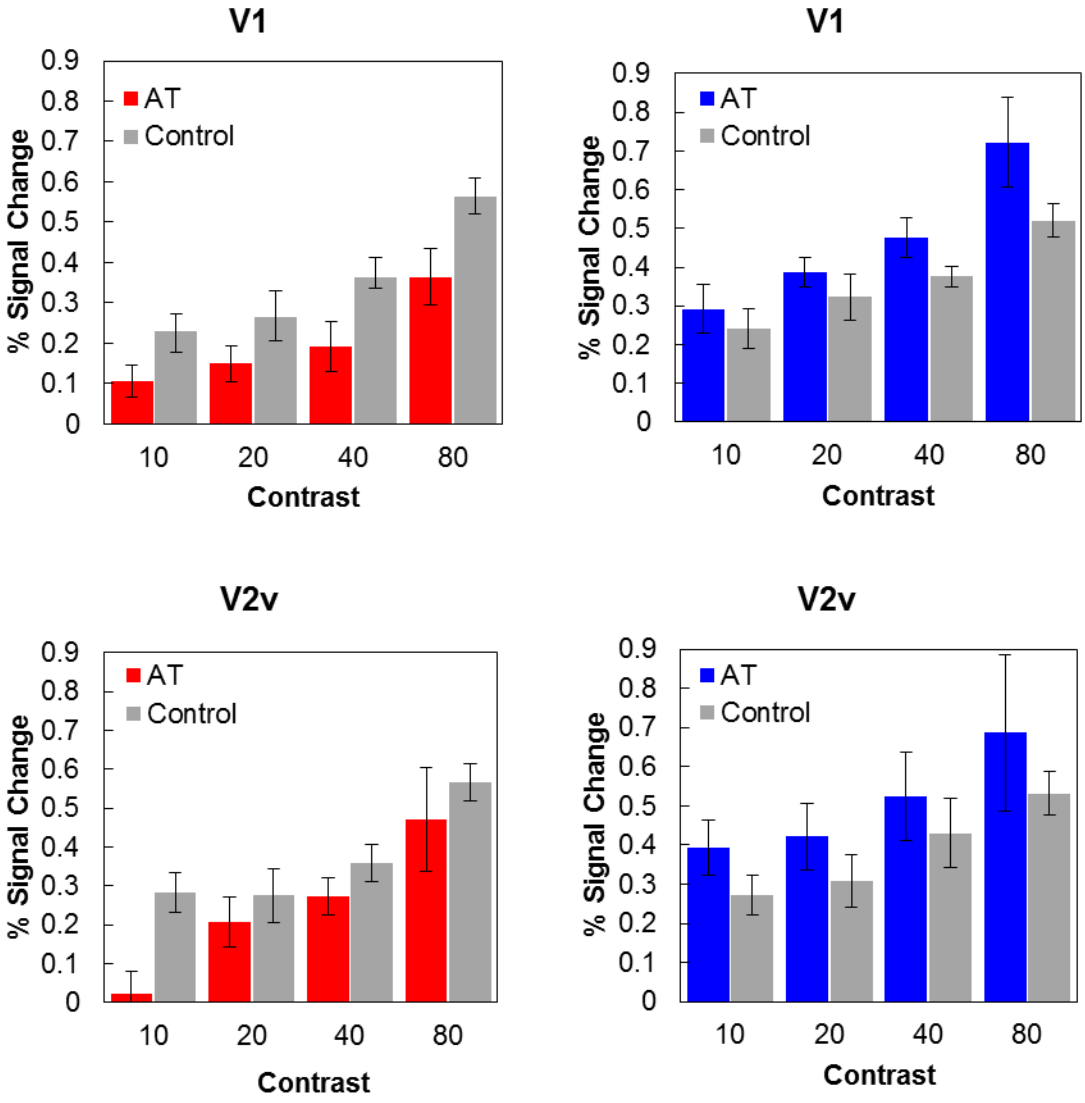

Figure 9. *Continued on next page.*

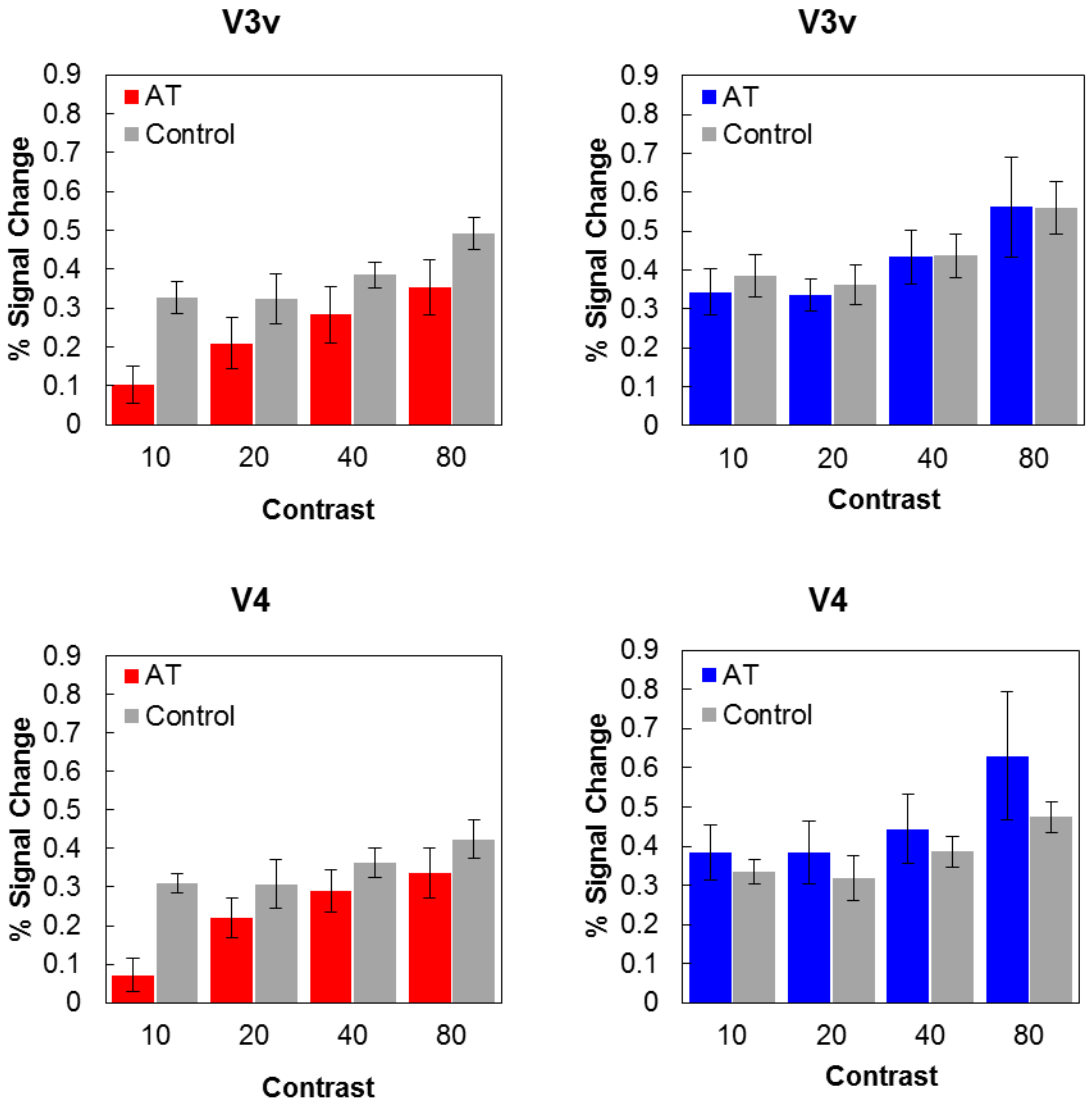

Figure 9. *The mean responses at each contrast level in percent signal change for Exp. 2, the high-attentional fixation task. Each row is activation in a region of interest, combined across the left and right hemispheres (V1, V2v, V3v, and V4), column 1 shows response to L/(L+M) stimuli, and column 2 shows activation to S/(L+M) stimuli.*

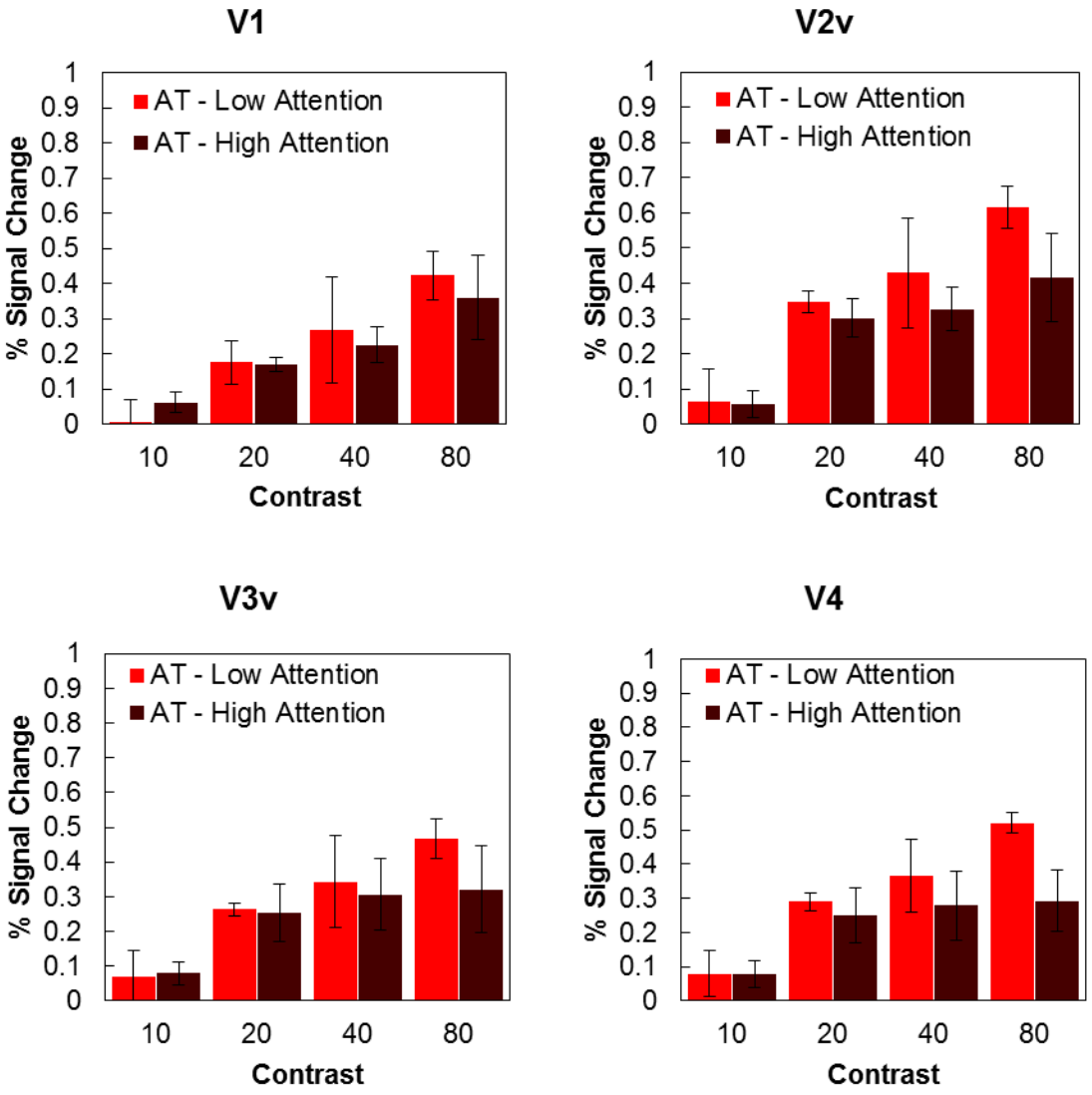

Figure 10. *Continued on next page.*

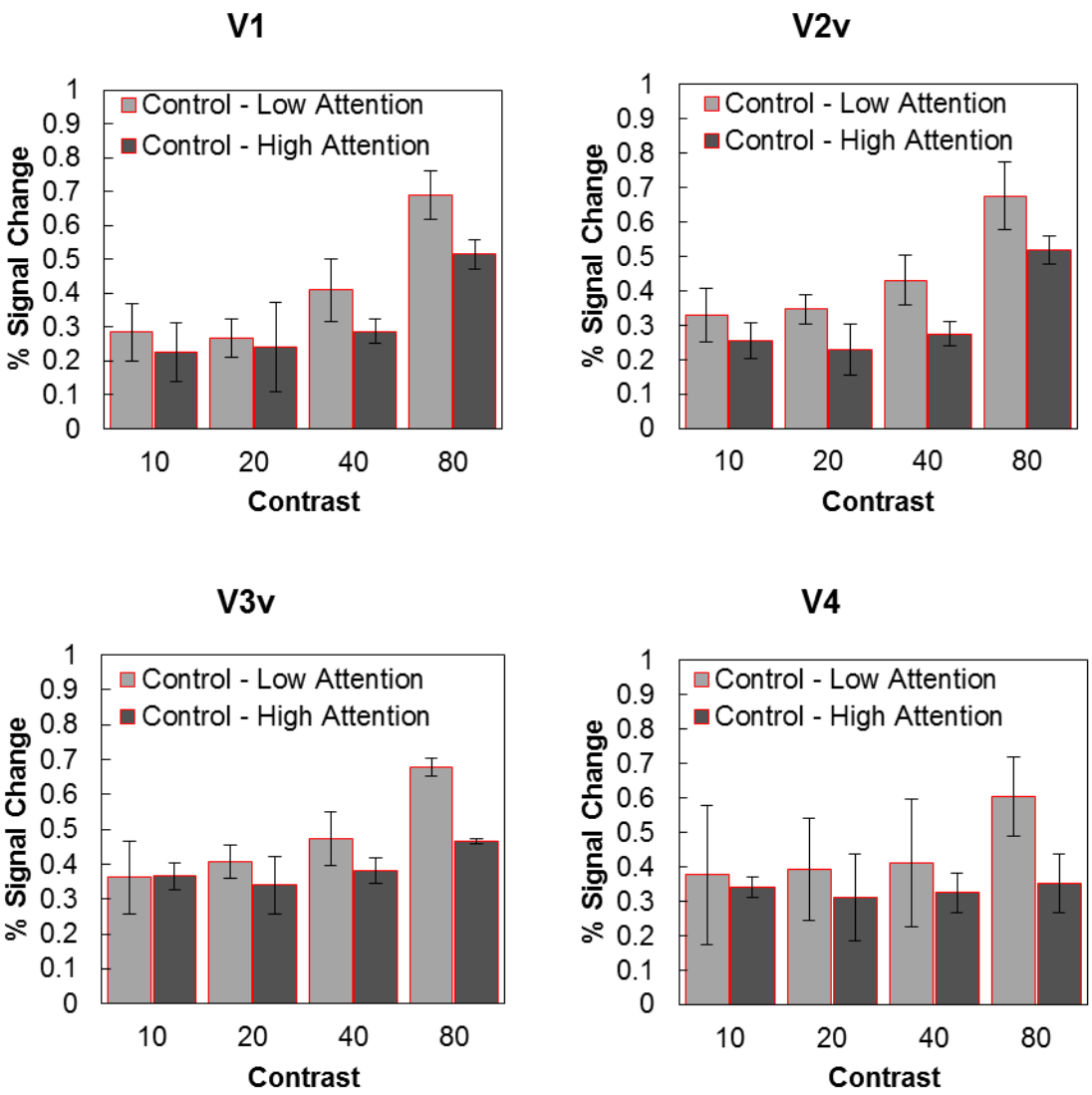

Figure 10. *The BOLD response to L-M stimuli for AT observers and controls, from Experiment 1 (low attention) and Experiment 2 (high attention) to L/(L+M) stimuli. Only observers who participated in both experiments are included here.*

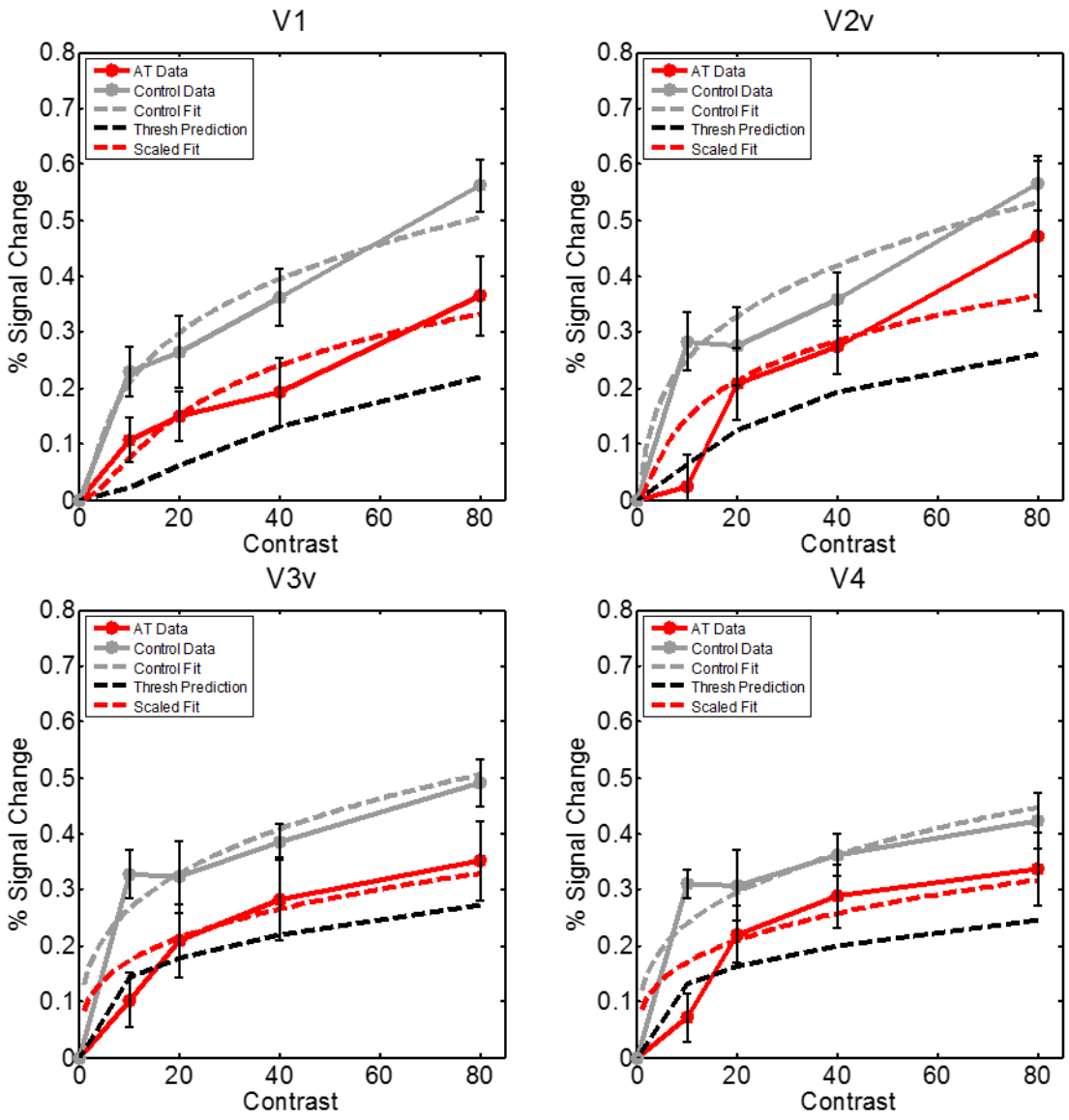

Figure 11. *Each plot shows the activation for a region of interest (V1-V4) for the L vs. M stimuli. The solid lines are the data, and the dashed lines are the fitted CRF functions. The gray is the control data, the red is the anomalous trichromat data, both averaged across observers. The black, dashed function is the CRF fitted to the control data, scaled using the threshold ratio of anomalous trichromats to controls.*

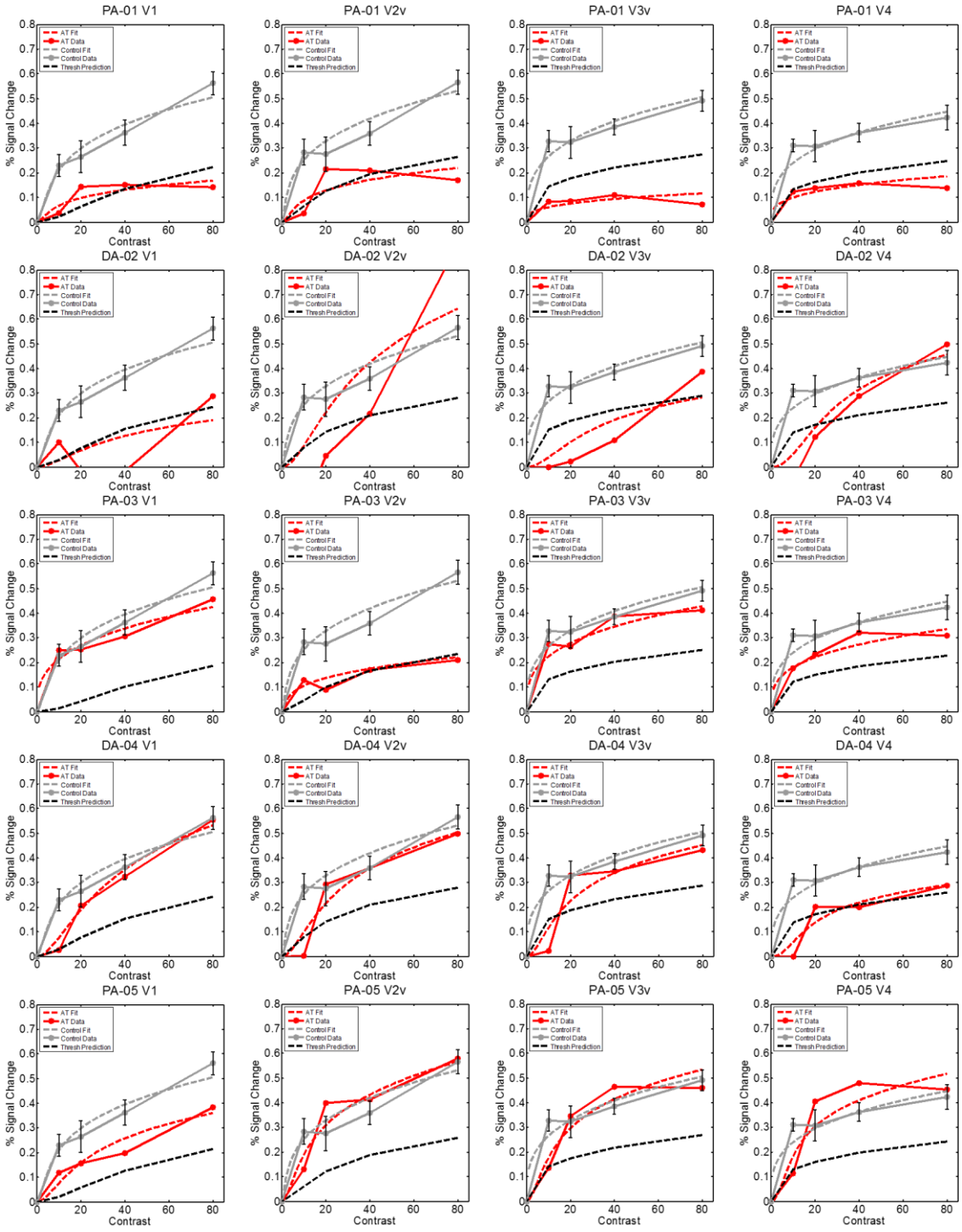

Figure 12. *Individual plots for each AT observer (row), and ROI (column). The AT observer's data is in red, and the control data is in gray. Solid lines represent the BOLD respones and dashed lines are the best fit CRF. The black line is the threshold prediction for each AT observer.*

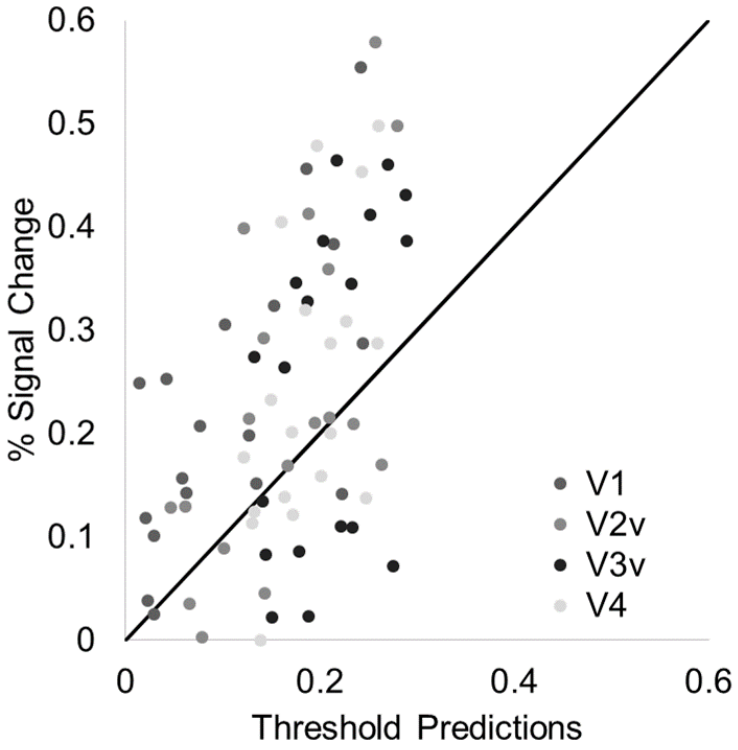

Figure 13. *Responses for each individual and ROI plotted against threshold predictions derived from the threshold ratio. Data falling above the solid line indicate responses where data was greater than predicted by the thresholds.*

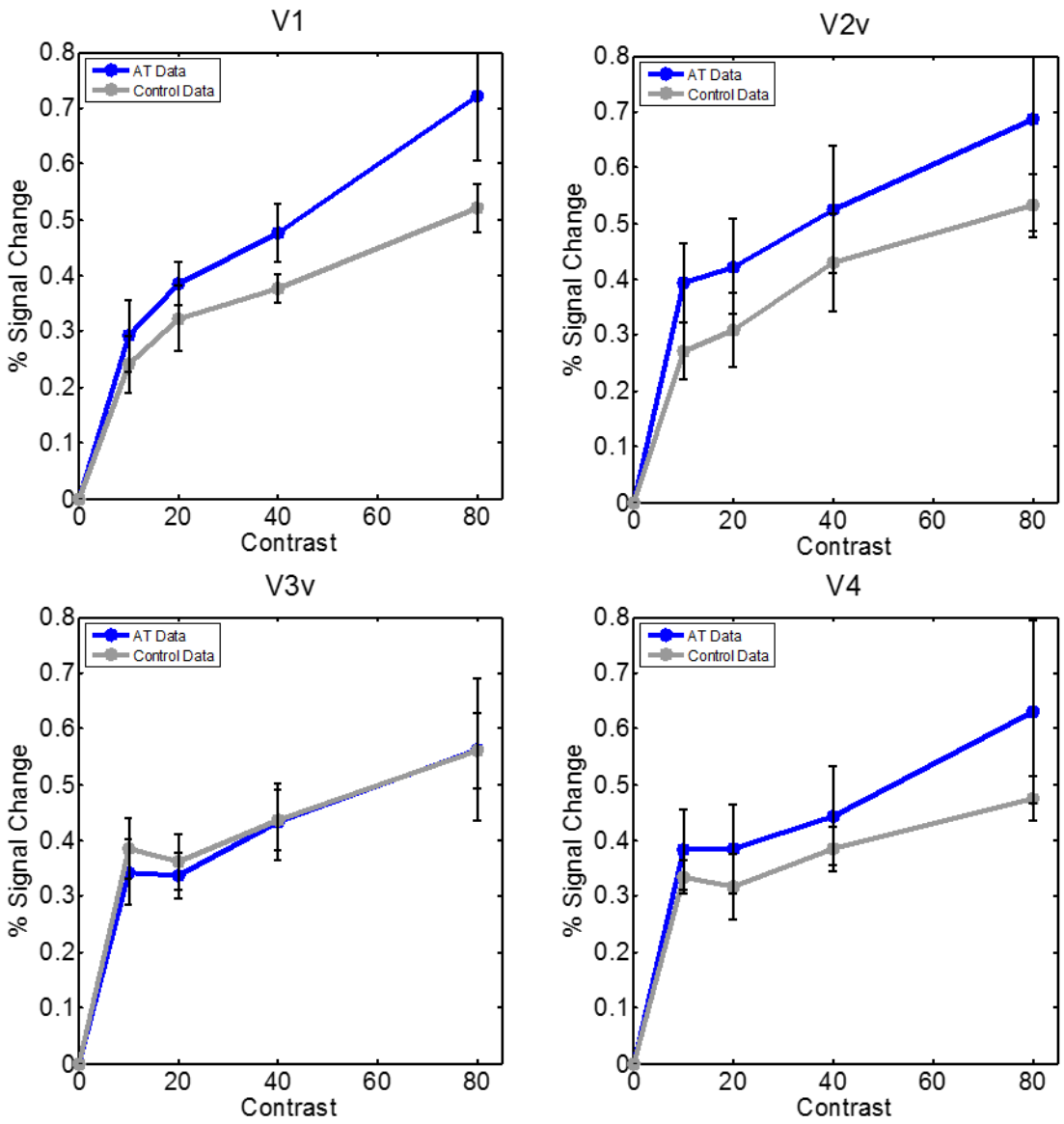

Figure 14. *Each plot shows the activation for a region of interest (V1-V4) for the S/(L+M) stimluli. The solid lines are the BOLD repsonses averaged across participants, and the dashed lines are the fitted CRF functions. The gray is the control data, the blue is the anomalous trichromat data, both average across observers.*

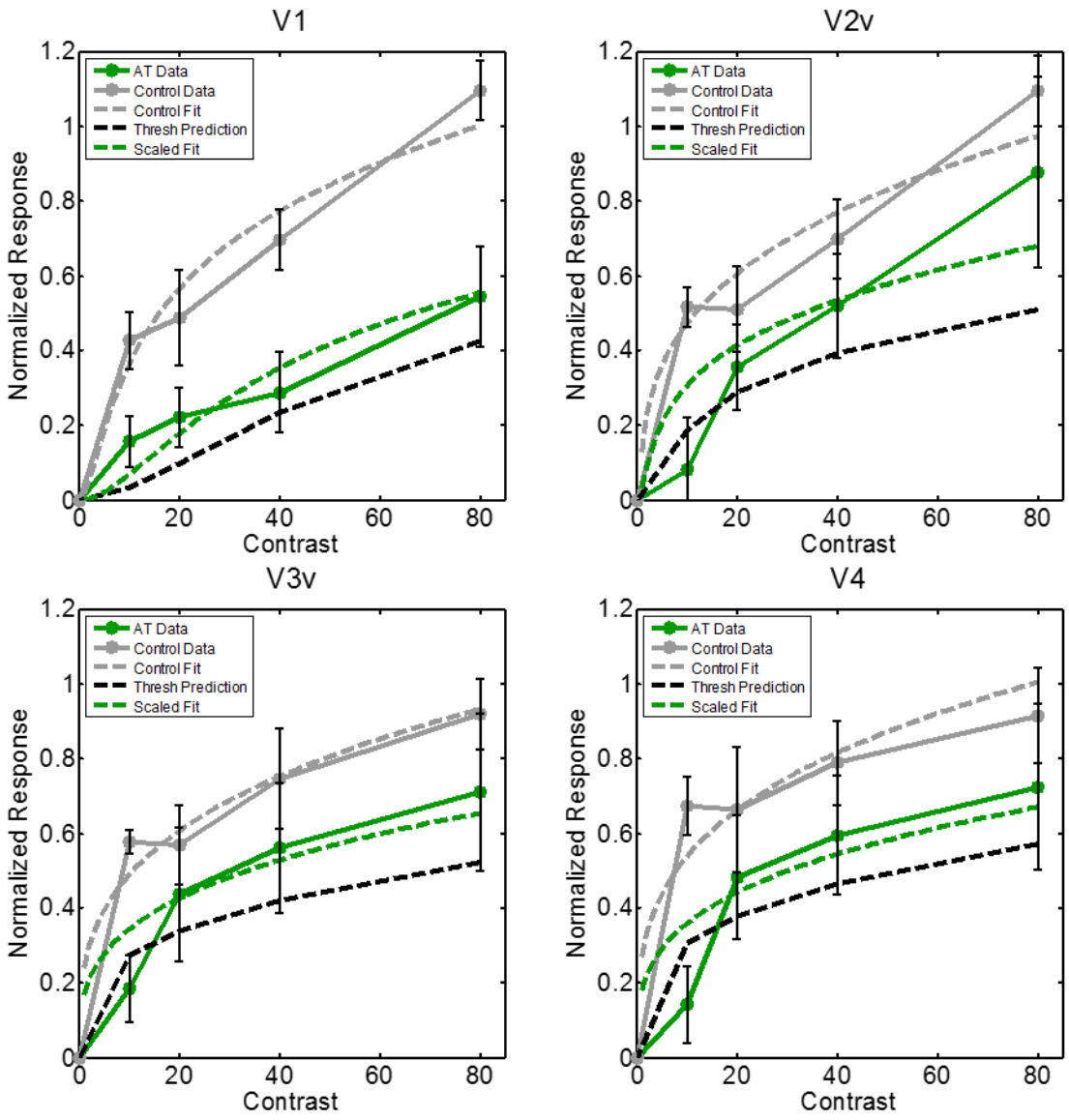

Figure 15. *Each plot shows results from a region of interest analysis (V1-V4) for the L/(L+M) data, normalized to the S/(L+M) data, where the response to S/(L+M) at contrast 80 is 1 and the baseline is 0. The solid lines are the actual data averaged across participants, and the dotted lines are the best fit CRF (green is AT, and gray is control). The black line was derived using the constants from the control data, scaled by the threshold ratio of AT to control.*

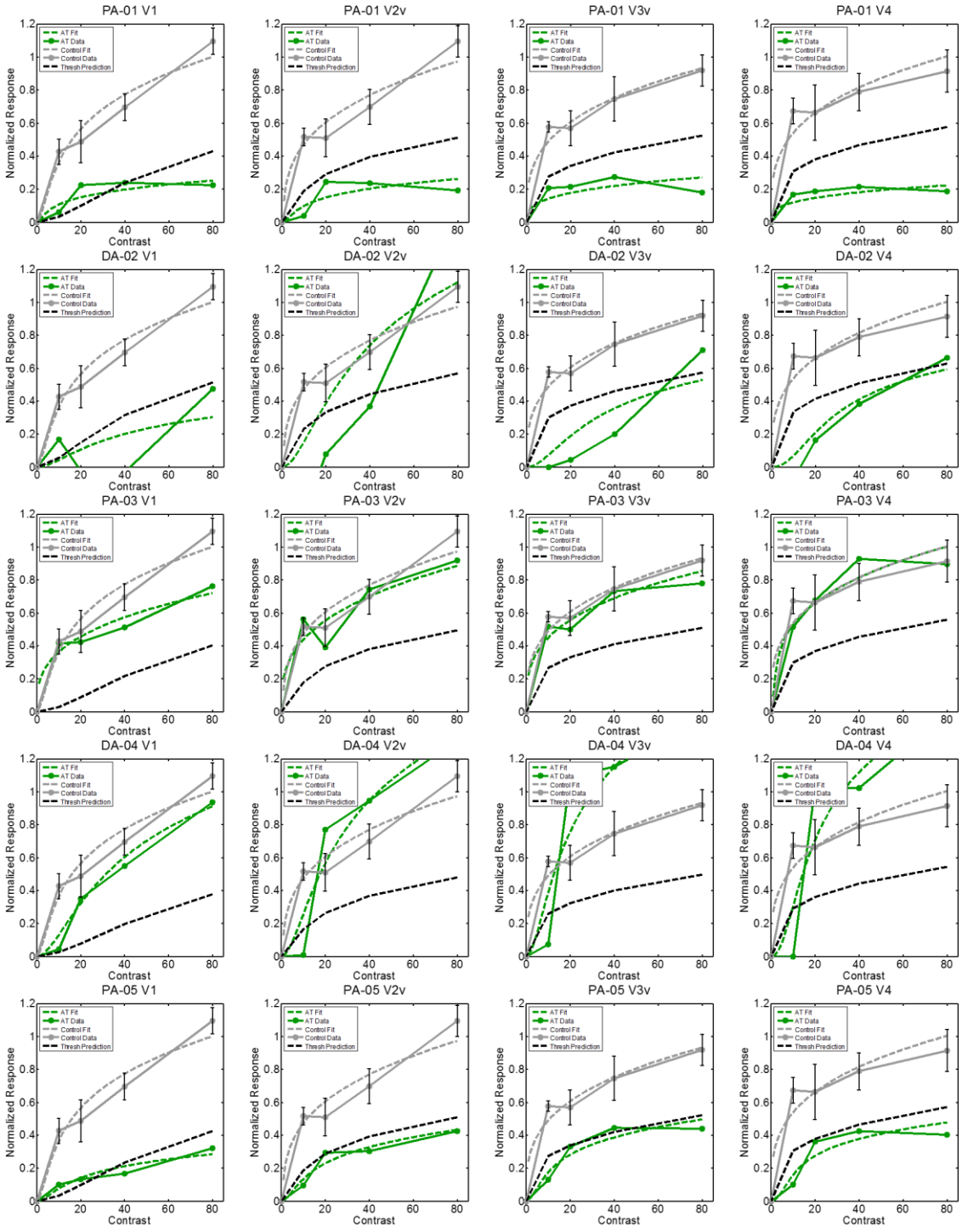

Figure 16. *The individual data, renormalized to each participants' S/(L+M), contrast 80 response. Rows are the individual data for each AT participant, columns are the regions of interest (V1-V4). "DA" indicates deuteranomalous and "PA" indicates protanomalous.*

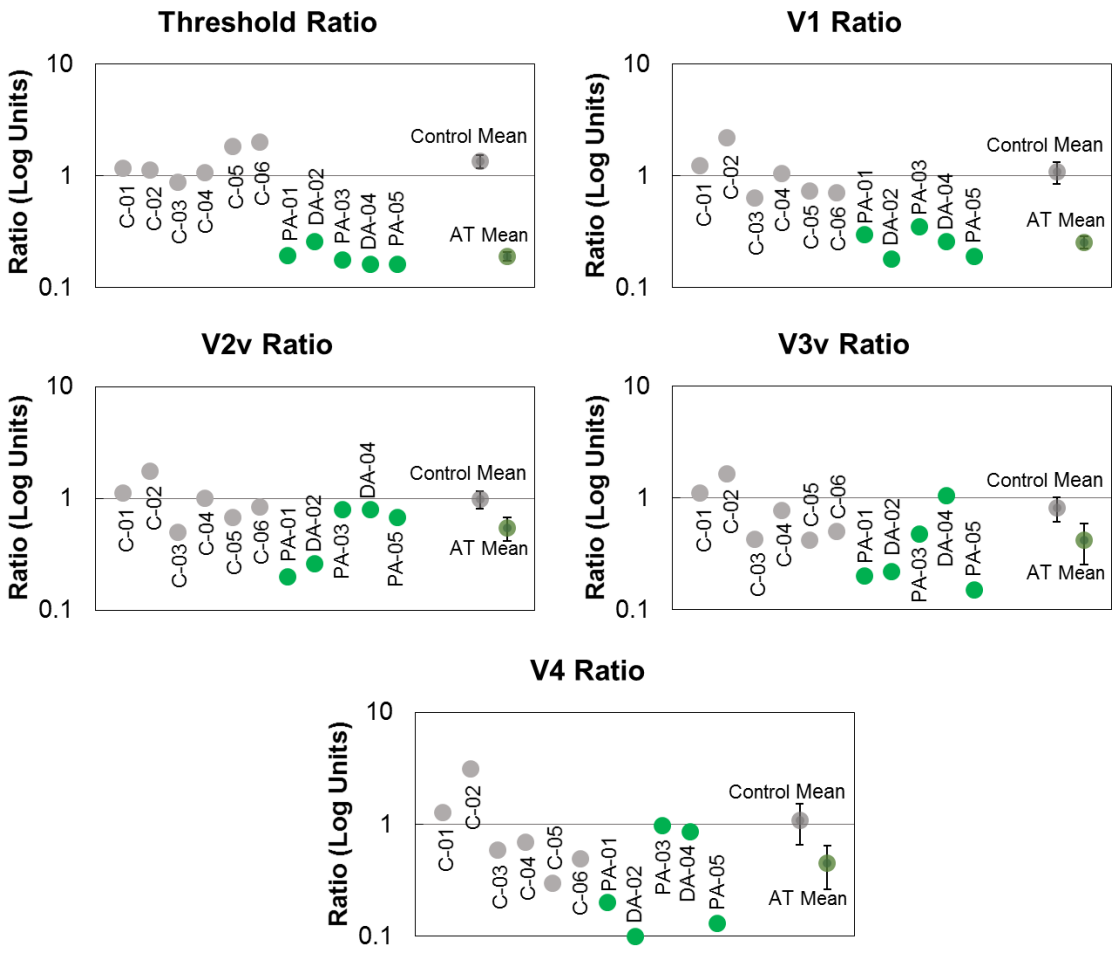

Figure 17. *Threshold ratios (L/(L+M) vs. S/(L+M)) and scaled MRI activation ratios (L/(L+M) vs. S/(L+M)) for individual observers (x-axis). A ratio of 1 would indicate equal responses for the L-M and S axes.*

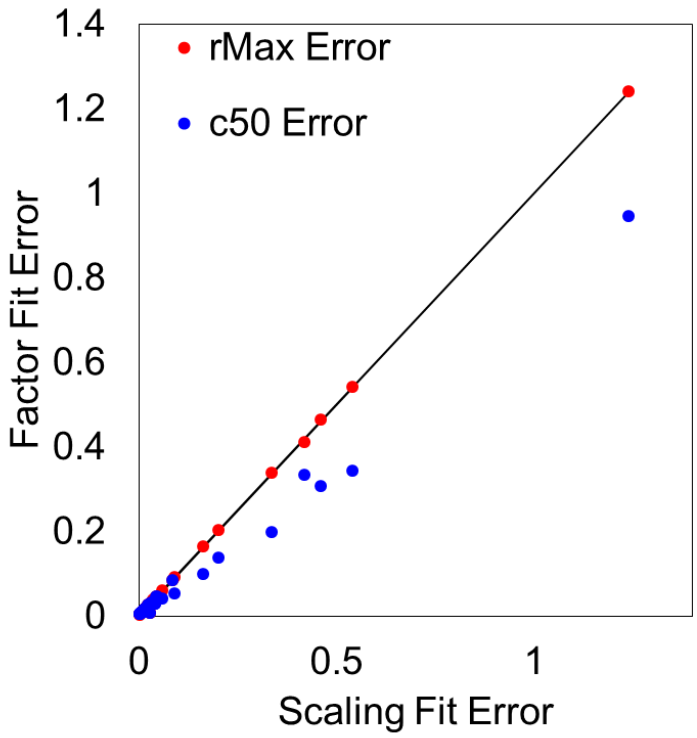

Figure 18. *SS error of the scaled fits from controls to AT observers, vs. the scaled fits derived from modulating either rMax (red) or c50 (blue).*

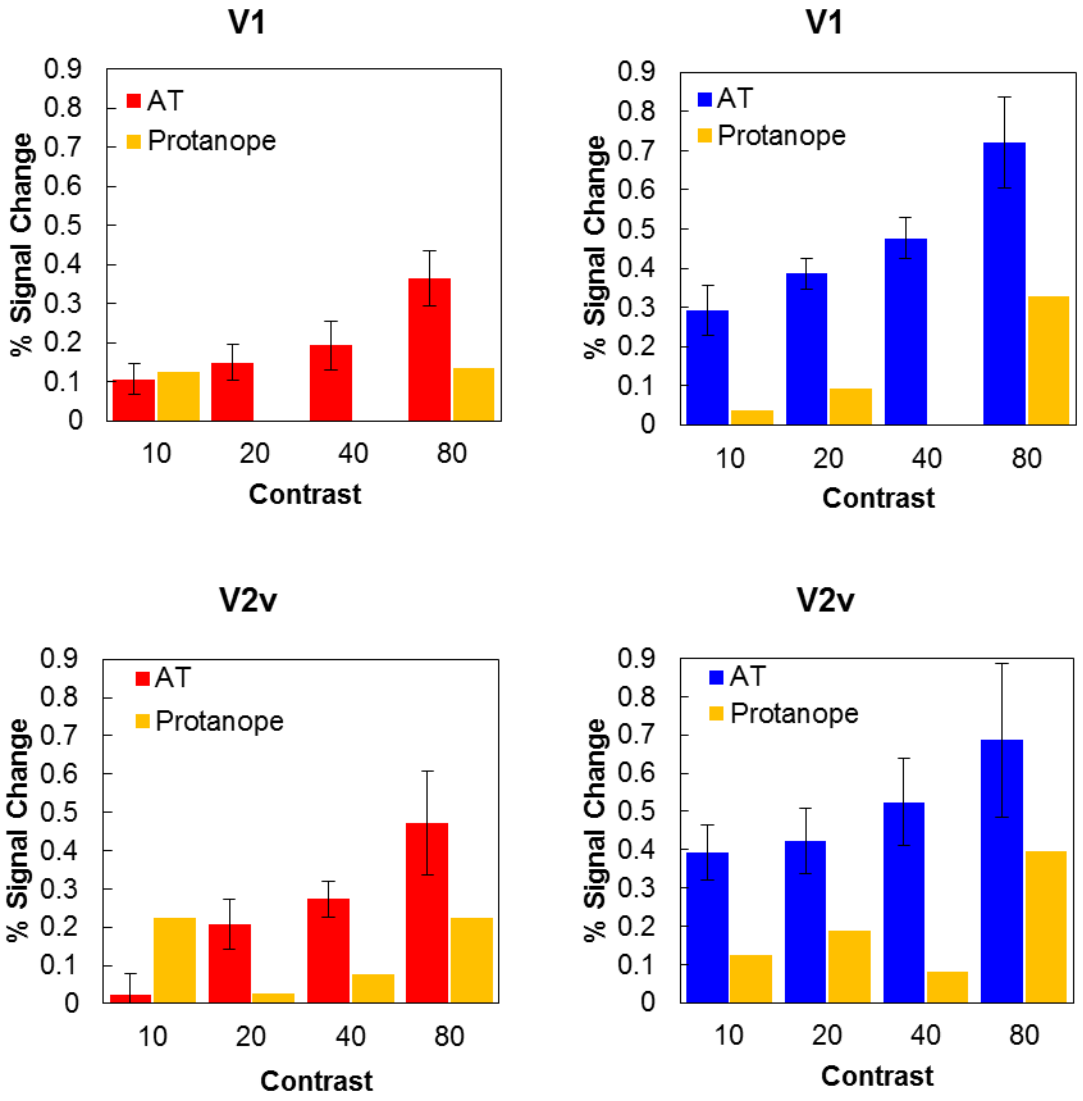

Figure 19. *Continued on next page.*

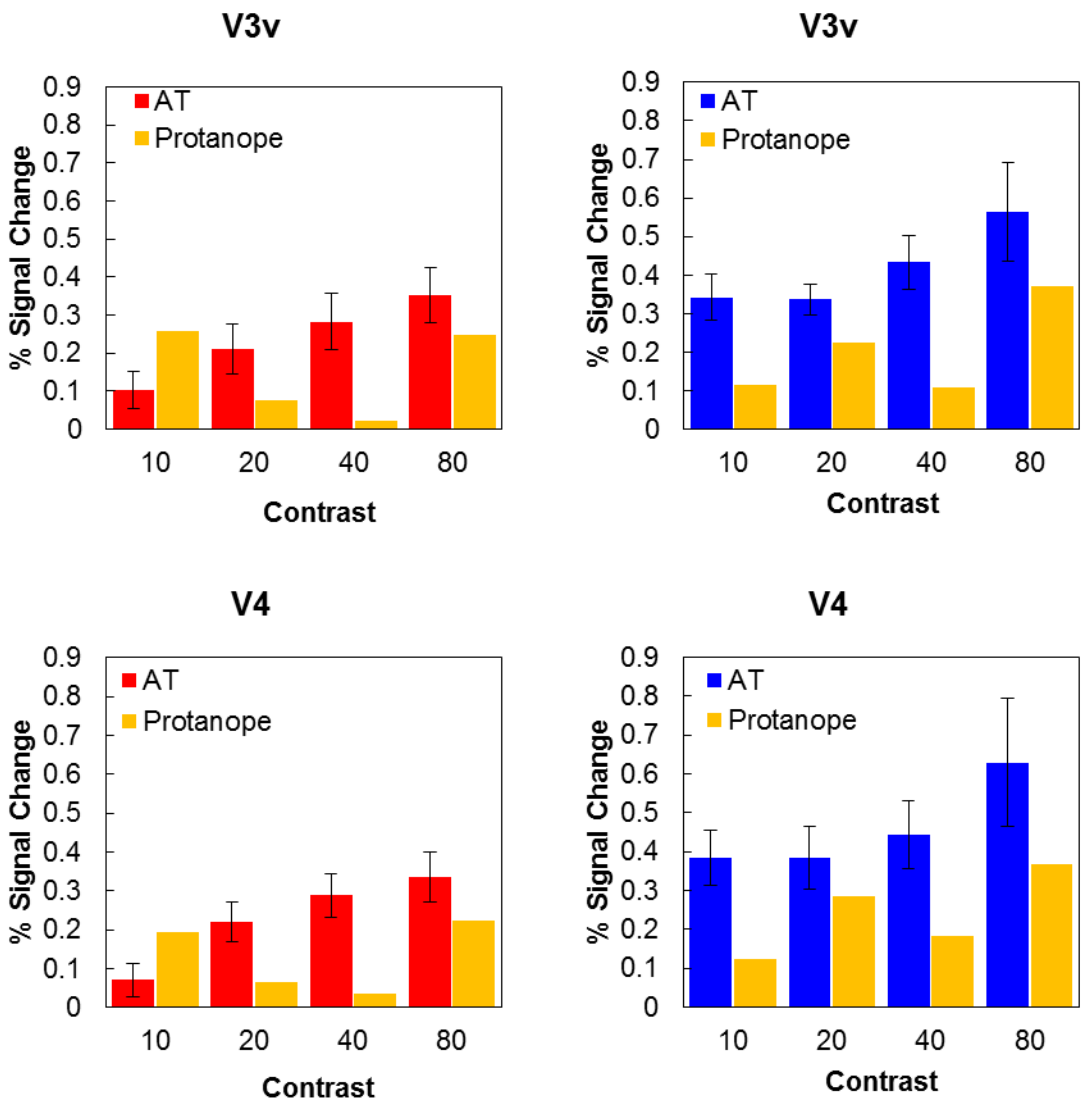

Figure 19*. Activation to chromatic contrast levels for one protanopic participant, as well as the mean activation across AT observers for areas V1, V2v, V3v, and V4. Absent bars indicate no significant activation above baseline. As in earlier graphs, the left column shows activation to the L/(L+M) stimuli, and the right column shows activation to the S/(L+M) stimuli.*

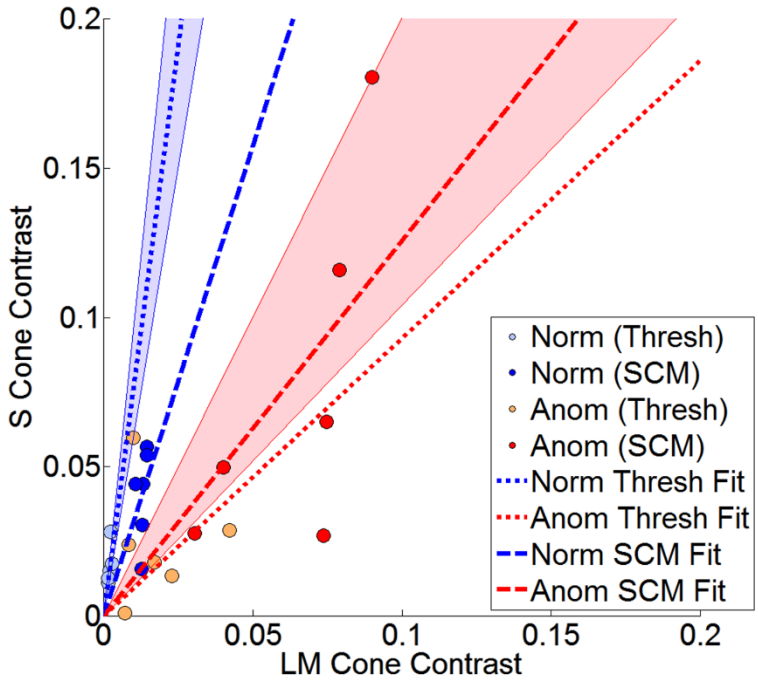

Figure 20. *Results from Vanston et al., 2017. Suprathreshold contrast matches were consistent with predictions based on AT thresholds.*