Functional magnetic resonance imaging adaptation reveals a noncategorical representation of hue in early visual cortex

Andrew S. Persichetti Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Sharon L. Thompson-Schill Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Omar H. Butt Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

David H. Brainard Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Geoffrey K. Aguirre Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

Color names divide the fine-grained gamut of color percepts into discrete categories. A categorical transition must occur somewhere between the initial encoding of the continuous spectrum of light by the cones and the verbal report of the name of a color stimulus. Here, we used a functional magnetic resonance imaging (fMRI) adaptation experiment to examine the representation of hue in the early visual cortex. Our stimuli varied in hue between blue and green. We found in the early visual areas (V1, V2/3, and hV4) a smoothly increasing recovery from adaptation with increasing hue distance between adjacent stimuli during both passive viewing (Experiment 1) and active categorization (Experiment 2). We examined the form of the adaptation effect and found no evidence that a categorical representation mediates the release from adaptation for stimuli that cross the blue–green color boundary. Examination of the direct effect of stimulus hue on the fMRI response did, however, reveal an enhanced response to stimuli near the blue–green category border. This was largest in hV4 and when subjects were engaged in active categorization of the stimulus hue. In contrast with a recent report from another laboratory (Bird, Berens, Horner, & Franklin, 2014), we found no evidence for a categorical representation of color in the middle frontal gyrus. A post hoc whole-brain analysis, however, revealed several regions in the frontal cortex with a categorical effect in the adaptation response. Overall, our results support the idea that the representation of color in the early visual cortex is primarily fine grained and does not reflect color categories.

Introduction

Color is the perceptual correlate of continuous physical variables. While people can make fine discriminations between stimuli of different wavelength composition, we also routinely divide colors into a relatively small number of discrete categories and find it natural to use basic color terms to name stimuli in each of these categories (Berlin & Kay, 1969). A categorical transition must occur somewhere between the initial encoding of the continuous spectrum of light by the cones and the verbal report of the name of a color stimulus. In this article, we study the representation of color in human visual cortex and ask whether there is evidence for categorical coding at this relatively early stage of visual processing.

Evidence from recent neuroimaging studies suggests that a categorical representation of color does not have its roots in primary visual cortex but rather may emerge in extrastriate cortex. Brouwer and Heeger (2013) demonstrated that multivoxel patterns in human visual areas hV4 and VO1 reveal a categorical clustering of 12 discrete colors when subjects actively categorized

Citation: Persichetti, A. S., Thompson-Schill, S. L., Butt, O. H., Brainard, D. H., & Aguirre, G. K. (2015). Functional magnetic resonance imaging adaptation reveals a noncategorical representation of hue in early visual cortex. *Journal of Vision, 15*(6):18, 1–19, http://www.journalofvision.org/content/15/6/18, doi:10.1167/15.6.18.

doi: 10.1167/15.6.18 Received January 1, 2015; published May 22, 2015 ISSN 1534-7362

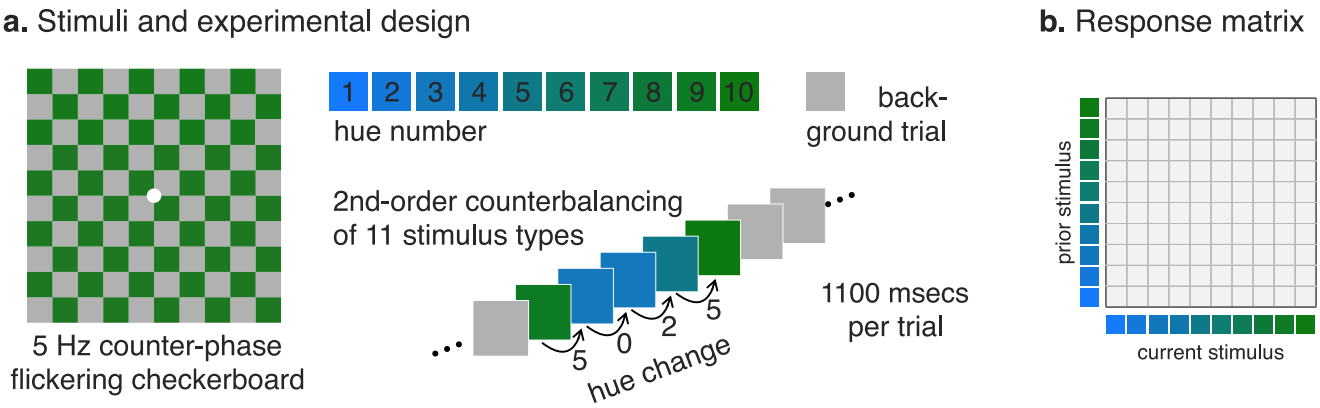

Figure 1. Stimuli and experimental design. (a) There were 10 stimulus colors, equally spaced from blue to green. For convenience, we use a 1-to-10 numerical hue scale to describe the stimuli throughout, with each step on this scale equal to a constant change in Munsell hue. Each of the 10 colors was displayed as a counterphase flickering (5 Hz) checkerboard contrasting the color with a neutral background. In Experiment 1, subjects attended to flickering of the central fixation dot. In Experiment 2, subjects actively categorized the color of the stimuli. The stimulus presentation sequence was counterbalanced to the second order. (b) The first-order response matrix codes each trial based on the stimulus hue of the current trial and the hue of the trial immediately prior.

stimuli but that this is not the case in earlier visual areas. Bird, Berens, Horner, and Franklin (2014) measured blood oxygen level dependent (BOLD) functional magnetic resonance imaging (fMRI) adaptation and reported a larger BOLD response for cross-category compared with within-category sequential color changes in the middle frontal gyrus (MFG) but not in visual cortex. The interpretation of the result of Bird et al. (2014) is complicated, however, as the stimuli themselves did not evoke a significant increase in mean BOLD signal within visual cortex. Similarly, studies of event-related potentials (ERP) find a larger late mismatch negativity response for between-categories compared with within-category hue pairs. This was found for both established (He, Witzel, Forder, Clifford, & Franklin, 2014) and newly learned (Clifford et al., 2012) color categories and suggests that a categorical representation emerges at a later stage.

Here, we examine the neural representation of color in primary and extrastriate visual cortex using an fMRI adaptation design. We examined whether there is a transition from a fine-grained toward a categorical representation of color along the ventral visual pathway from V1 to hV4. We presented 10 stimuli that varied in small steps along a green–blue continuum in a continuous carryover fMRI experiment, which allowed us to simultaneously measure both the direct effect of stimulus hue and the sequential effect of hue change (Aguirre, 2007). The color stimuli were presented as a flickering checkerboard designed to produce a robust response in the visual cortex. We studied the release from adaptation as a function of hue change and examined whether this was modulated by a category transition. Although our stimuli differed, our experimental logic shares much with that of Bird et al. (2014). Therefore, we also examined the representation of color within their bilateral MFG region in an attempt to replicate their findings.

Experiment 1

Method

Subjects

Ten subjects between the ages of 21 and 28 years (six females, four males) with normal or corrected-to-normal vision were recruited from the Philadelphia area. All subjects had normal color vision as verified by administration of the Ishihara Test for Color Deficiency (Ishihara, 1936). Two of the subjects were members of the Aguirre laboratory, while the other eight were naïve to the purposes of the experiment. Informed consent was obtained from all subjects. All procedures were approved by the Institutional Review Board of the University of Pennsylvania and adhered to the tenets of the Declaration of Helsinki.

Stimuli

Each of 10 color stimuli was a counterphase flickering (5 Hz) checkerboard contrasting the color and a neutral background (Figure 1a). The 10 colors used in the checkerboards were computer renditions of Munsell papers with a Munsell value of 5.5 and Munsell hue values that ranged from 3.89 B to 3.89 G in equally spaced hue angle steps. For each rendered paper, the chroma was maximized within the gamut of the display device and the stimuli were collectively

Stimulus	Munsell renotation	xyY coordinates (cd/m^2)
Background	Neutral (N 7/)	0.310, 0.316, 807.5
Test hue 1	3.890B 5.5/7.7	0.206, 0.267, 467.0
Test hue 2	1.670B 5.5/7.4	0.210, 0.284, 467.0
Test hue 3	9.44BG 5.5/7.2	0.217, 0.298, 467.0
Test hue 4	7.22BG 5.5/7.1	0.220, 0.313, 467.0
Test hue 5	5.00BG 5.5/7.2	0.225, 0.329, 467.0
Test hue 6	2.78BG 5.5/7.5	0.231, 0.349, 467.0
Test hue 7	0.56BG 5.5/7.7	0.237, 0.365, 467.0
Test hue 8	8.330G 5.5/8.1	0.242, 0.383, 467.0
Test hue 9	6.110G 5.5/8.5	0.247, 0.401, 467.0
Test hue 10	3.890G 5.5/9.0	0.258, 0.426, 467.0

Table 1. Munsell paper values and corresponding Commission Internationale de l'Eclairage (CIE) chromaticity and luminance (cd/m^2) of the 10 colored stimuli and background gray.

scaled to lie within the overall luminance range that could be displayed. The background was a computer rendition of a Munsell N 7/neutral paper. The colorimetric specifications of the 10 colors and the background are provided in Table 1. For convenience, we use a 1-to-10 numerical hue scale to describe the stimuli throughout, with each step on this scale equal to a constant change in Munsell hue.

The checkerboards themselves consisted of a 10×10 grid of square checks and subtended 10° of visual angle (1° per check). Each checkerboard had a small white fixation dot, subtending 0.15°, at its center. The modulation of each check was one sided so that the check alternated between the gray background and the color specified for that checkerboard.

Procedure (fMRI)

Given the 10 possible colored checkerboards and a null (uniform gray field) stimulus, a de Bruijn sequence ($k = 11$, $n = 3$) was created that counterbalanced stimulus presentation to the second level. A path-guided sequence was crafted that ordered sequential transition sizes in time such that experimental power was optimized to detect a linear effect of neural adaptation following convolution with the BOLD hemodynamic response function (Aguirre, Mattar, & Magis-Weinberg, 2011). Second-level counterbalance is required to allow valid inferences to be drawn regarding first-level (one-back) effects (Aguirre et al., 2011, appendix A). The initial length of the sequence was 1,331 trials (11^3). The length of null trials was doubled and in some cases tripled to increase statistical power for detection of main stimulus effects versus the null (Aguirre, 2007, appendix A), which increased the total number of trials to 1,500. The sequence was then broken into five sections of 300 elements each.

At the start of a scan, several seconds are required for the BOLD response to build to a steady state, and at the end of a scan, the hemodynamic response to the last presented stimulus continues after data collection has ceased. These effects were accounted for by having each subsequence repeat the last 30 elements from the prior subsequence at the onset (taken circularly, so that the last 30 elements from the fifth sequence were prepended to the first). Each subsequence was run during a separate BOLD fMRI scan of 330 trials (1100 ms each), corresponding to 121 trials (3000 ms each). In analysis, the data from the first 11 trials of each run were discarded, and the resulting data, consisting of 110 trials each, were concatenated and modeled as if the complete sequence had been run in a single scan.

The experiment comprised 10 scans that were made up of the five sequences described above followed by those same five sequences in reverse order. These data were collected over two separate scanning sessions. On each trial, a checkerboard stimulus was displayed for 1000 ms followed by a blank screen for 100 ms. The fixation dot blinked to black for 250 ms on a random schedule (~100 times per scan), and participants were instructed to indicate a blink event by pressing a button.

Scanning

We collected echoplanar BOLD fMRI data on a 3 Tesla Siemens Trio (Siemens Medical Solutions USA, Inc., Malvern, PA) using an eight-channel array coil at a repetition time (TR) of 3 s, with 3-mm isotropic voxels covering the entire brain. A high-resolution anatomical image (3-dimensional Magnetization Prepared Rapid Acquisition GRE, MPRAGE; 160 slices, 1-mm isotropic voxels, TR = 1.62 s, echo time = 3.09 ms, inversion time = 950 ms, field of view = 250 mm, flip angle = 15°) was acquired for each subject. The anatomical scan was performed at the end of the session, after the functional scans.

Analysis of anatomical images and definition of regions of interest (ROIs)

Anatomical data from the subjects were processed using the FMRIB Software Library toolkit (http://www.fmrib.ox.ac.uk/fsl/) to correct for spatial inhomogeneity and to perform nonlinear noise reduction. Brain surfaces were reconstructed and inflated from the MPRAGE images using the FreeSurfer (Version 5.1) toolkit (http://surfer.nmr.mgh.harvard.edu/) as described previously (Dale, Fischl, & Sereno, 1999; Fischl & Dale, 2000). The anatomical image from each subject was registered to the fsaverage-sym template space (Fischl, Sereno, & Dale, 1999; Greve et al., 2013).

Visual areas were defined for each subject using a cortical surface template of retinotopic organization (Benson, Butt, Brainard, & Aguirre, 2014). V1, V2, and

V3 regions extended from 0° to 10° eccentricity. The hV4 region was defined as lying between the outer border of ventral V3 (extending to 10° eccentricity) and the posterior transverse collateral sulcus (Witthoft et al., 2014). These regions were defined within the pseudohemisphere fsaverage-sym surface template space and were projected back to the volumetric anatomy of each subject for analysis. As no meaningfully different responses were observed between visual areas V2 and V3, the data from these regions were combined.

An additional ROI was defined in the bilateral MFG. The regions were those reported by Bird et al. (2014) and were kindly supplied by the authors as a defined mask within the Montreal Neurological Institute brain atlas space. To create ROIs for use in the current study, we registered the volumetric anatomical data from each subject to the Montreal Neurological Institute space using SPM8 (www.fil.ion.ucl.ac.uk/spm). The MFG regions of Bird et al. (2014) were then projected to the individual volumetric anatomical space of each of our subjects.

Analysis of fMRI data

After image reconstruction, the data were sinc interpolated in time to correct for the fMRI slice-acquisition sequence and motion corrected. Minimal spatial smoothing was performed with a 1.5-mm full-width half-max three-dimensional Gaussian kernel. The echoplanar data from each subject were coregistered to subject-specific anatomy in FreeSurfer using FSL-FLIRT with 6 *df* under a FreeSurfer wrapper (bbregister). Within-subject statistical models were created using the modified general linear model (Worsley & Friston, 1995). Experimental conditions were modeled as described below and convolved with an average hemodynamic response function. Nuisance covariates included effects of scan and global signals. Large transients in the data caused by head movements were identified as time points in the raw data with a sequential transition in signal that exceeded 2 *SD* from the mean of the time series. These time points were modeled as a delta function. There were on average 10 (range: 3–20) transients modeled per subject.

A set of reference-coded covariates modeled the effect of the color stimuli within the fMRI data. Construction of the covariates began with the coding of each stimulus as occupying a position within a first-order response matrix (Figure 1b). This matrix codes each trial based on the stimulus hue of the current trial and the hue of the trial immediately prior. Trials on which a hue is presented and the prior trial was a null stimulus are excluded from this scheme and coded instead as new trials. In all analyses, a separate (nuisance) covariate modeled the effect of the new trials versus the null trials and was not further evaluated.

We used two approaches to analyze the data, each based on a set of covariates that modeled the response matrix. The first approach used a set of 100 covariates, with each covariate modeling as a delta function the neural response to those trials on which a particular transition between a pair of hues occurred. Such a model is complete but risks overfitting the data and is not readily interpretable. The explanatory power of this model was assessed and used as a point of comparison.

The second approach is the one adopted for the bulk of the results presented here and is referred to below as the 10-covariate polynomial model. The entire response matrix was modeled with a limited polynomial basis set that targeted direct and sequential effects (Aguirre, 2007). The direct effect of the hue currently presented in a trial was modeled as a neural response that was a function of the hue (1–10 on our hue scale) of the currently presented stimulus. A polynomial expansion of this relationship was created, such that separate covariates modeled a linear relationship between hue position and neural response, a quadratic relationship between hue position and neural response, and so on. Each higher level polynomial was rendered orthogonal with respect to the set of lower level covariates. This orthogonalization, while of no consequence for the overall ability of the model to account for variance in the data, facilitated interpretation of the measured effects. Empirically, a fourth-order polynomial set was found to be sufficient to model the data.

The model also contained a complementary set of polynomial covariates that modeled the modulatory effect of a change in hue from the prior to the current stimulus. These covariates capture effects described as neural adaptation. The linear adaptation covariate modeled neural response for the current trial as being proportional to the change in hue number from the prior trial. Quadratic, cubic, and quartic adaptation relationships were also modeled, with each higher level polynomial rendered orthogonal with respect to the lower order components.

An additional covariate modeled the possibility of a differential neural adaptation effect for stimulus transitions of three hue steps that occurred in the center of the stimulus space (and thus crossed the perceptual categorical boundary) versus transitions of the same size that occurred at the periphery of the stimulus space.

Each of the covariates described above was mean centered, and a separate main effect covariate was included to model the mean response to all trials within the response matrix compared with the null trials. The complete model therefore contained 10 covariates: one main effect, four polynomial components that modeled the direct effect of the current hue, four polynomial

components that modeled the neural adaptation effect of sequential hue change, and one covariate that modeled an enhanced neural response to categorical transitions. The counterbalanced trial sequence and the set of covariates used (prior to and following convolution with the hemodynamic response function) are available for download (https://cfn.upenn.edu/aguirre/wiki/public:public_stimuli_and_data).

Covariate time shift

The set of covariates was convolved by a hemodynamic response function to provide predictors for the BOLD fMRI data. As differences in the time to peak across subjects can degrade the ability of the covariates to model stimulus responses (Aguirre, Zarahn, & D'esposito, 1998), covariates were time shifted for each subject to best fit the data in area V1. For each subject, the average across-voxel fMRI signal was obtained from the V1 region of interest. The correlation of this signal with only the main effect covariate was measured with sinc interpolated phase shifting of the covariate between $\pm$2000 ms in 250-ms increments. The time shift that yielded the highest correlation was retained for each subject and then applied to the entire set of covariates. Across subjects, temporal shifts ranged from $-$250 to 1625 ms; six of the 10 subjects had temporal shifts with an absolute value of less than 500 ms.

Interrogation of regions of interest

The primary analyses of this study were conducted by measuring the BOLD fMRI response across subjects within defined regions of interest. For each subject, within each ROI, those voxels that had a positive t value of 3.5 or greater for the main effect of stimulation were identified. The average fMRI signal within this set of responsive voxels within a region was obtained and then supplied to the modified general linear model. As no differences were observed between hemispheres, all data were averaged between hemispheres. The beta weight on each covariate was retained. The set of beta weights across subjects for a particular covariate were then averaged for plotting or subjected to t tests versus zero.

Thresholding by the main effect was not performed for those analyses in which the main effect itself was being measured and was never performed within the MFG ROI as no voxels responsive to the main effect were found. Instead, the average signal within the entire region identified by Bird et al. (2014) was obtained.

Creation of group maps

Whole-brain statistical maps were created for display of effects but were not used for statistical inference. For the display of any given covariate, the volumetric beta-weight map of that covariate was obtained for each subject and transformed to the fsaverage-sym space. The beta maps from the left and right hemispheres were averaged within a pseudoleft hemisphere within each subject and then smoothed with a 5-mm full-width at half-maximum, two-dimensional Gaussian kernel. The across-subjects statistical effect of the covariate was assessed at each vertex within the surface-space template against zero and expressed as a z value. The across-subjects z-map was then thresholded at $z = 2.5$, corresponding to a voxel-wise $p = 0.01$. A sign permutation test (Nichols & Holmes, 2002) was then used to assign an extent threshold that provides a map-wise, cluster-corrected $p = 0.05$ using the routines integrated in the FreeSurfer software package (Hagler, Saygin, & Sereno, 2006).

Psychophysics

During an additional behavioral experiment, we measured blue–green color naming. Subjects viewed the same flickering checkerboard stimuli that were used during scanning and were asked to make a button response corresponding to whether they judged the color to be blue or green. Trials advanced at the pace of subject responses, and response times were collected. All subjects completed this color-naming task while still in the scanner but after their functional scans had been completed (i.e., at the end of their second scanning session). The psychophysical measurements were made in the scanner to ensure that identical stimuli were used for both psychophysics and fMRI. They were made after functional scanning to ensure that the color-naming task did not prime neural responses to color during the passive-viewing fMRI task. For one subject, however, we discovered after the scanning sessions that the functional data from his first scan were not usable because of loss of synchronization between the scanner and the stimulus display computer. Naming data from this subject's unusable pair of scanning sessions were discarded, and this subject repeated all scans and reperformed the color-naming task 55 days later.

Display calibration

Stimuli were presented using a Sanyo SXGA projector (4200 lumens; SANYO Electric Co., Ltd., Osaka, Japan) with a Buhl long-throw lens (Buhl Industries, Fairfield, NJ) for rear projection onto a Mylar screen, which subjects viewed through a mirror mounted on the head coil in the scanner. The display system was calibrated in situ using a Photo Research PR-650 spectral radiometer (Photo Research, Inc., Chatsworth, CA). Spectral power distributions for each projector channel were measured, as were the channel

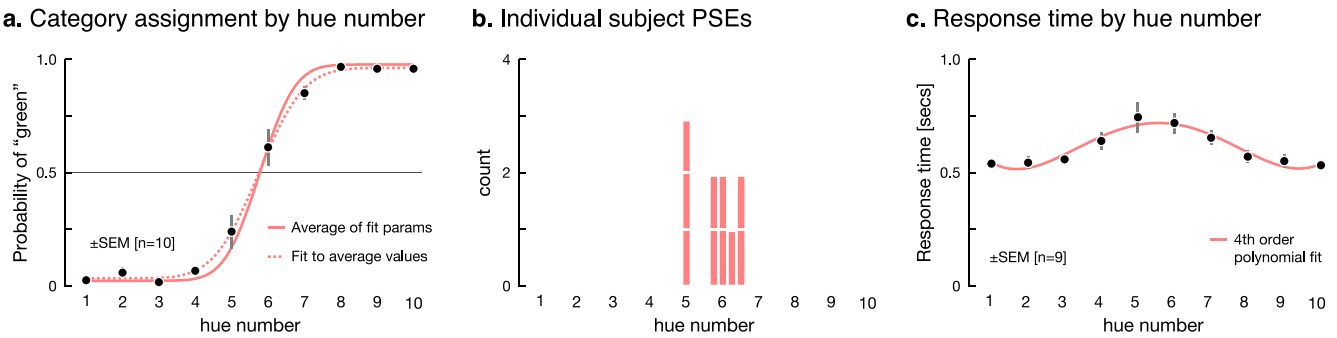

Figure 2. Psychophysical results in Experiment 1. (a) Blue–green categorization performance. Each point shows the average over subjects of the probability of a *green* response to each of our 10 colored checkerboards. The dashed red line is the fit of a cumulative normal to these average data. The solid red line is the cumulative normal obtained by averaging the mean and variance parameters of a cumulative normal fit to each subject's data. Error bars show ±1 standard error of the mean (*SEM*) and are smaller than the plotted points in some cases. (b) Histogram of PSEs obtained for individual subjects. (c) Response time data. Each point shows the average over subjects of the median within-subject response time required to categorize each colored checkerboard as blue or green. Fit is of a fourth-order polynomial. Error bars show ±1 *SEM*. Response time data were collected for nine of the 10 subjects.

input–output (gamma) functions. This information was used together with standard methods (Brainard, Pelli, & Robson, 2002) to determine the appropriate RGB settings to produce the desired tristimulus coordinates (Table 1) for each stimulus.

Results

Color categorization performance

Figure 2a shows the psychophysical blue–green categorization performance. Each point shows the average (across subjects) of the probability of a *green* response as a function of stimulus hue. The data are highly regular. The most extreme blue and green stimuli are reliably named *blue* and *green*, respectively, and there is a steady increase in the proportion of *green* responses that occurs between stimulus numbers 4 and 7. These data confirm that our stimulus set contains good exemplars of stimuli judged reliably as blue and green as well as exemplars of stimuli that lie near the blue–green category boundary.

To summarize the data, we fit a cumulative normal as a function of stimulus number to both the individual and average data, separately allowing for lapses for *blue* and *green* responses. The dashed line in Figure 2a shows the fit to the average data and indicates that the cumulative normal can provide a good parametric description of the responses. Note, however, that in the average data variation in the mean point of subjective equality (PSE) between subjects can produce a shallower aggregate slope than that for any of the individual subjects. For this reason, we chose to summarize the aggregate data by averaging the parameters of the cumulative normal fit to the data from each subject. The solid line in Figure 2a shows this fit. From this, we obtain an estimate of the blue–green category boundary along the 1-to-10 scale that we use for stimulus hue, taken as the stimulus value that yields an equal probability of *blue* and *green* responses according to the fit (solid red line). This value is 5.8. There was reasonable agreement between subjects in terms of the location of the blue–green category boundary (Figure 2b).

We also examined response times in the categorization task. These are highest for stimuli near the blue–green category boundary and fall off for less ambiguous stimuli (Figure 2c). This type of response time effect has long been observed in categorization experiments (Cross, Lane, & Sheppard, 1965).

Early visual cortex response to the stimuli

Before proceeding with our primary analyses of the fMRI data, we identified regions of the brain where responses were evoked by our stimuli. Both region-based and whole-brain analyses were conducted to measure the effect (across subjects) of the presentation of any of the 10 colored flickering checkerboards versus trials in which only the uniform gray background was present. Figure 3a shows this main effect of color trials across the brain (left and right hemisphere averaged). There was a highly significant increase (z scores > 9) in the BOLD signal in the occipital cortex. This response corresponds to the locations of visual areas V1, V2, V3, and hV4 at the eccentricities covered by our stimuli, as identified by our surface-based retinotopic template (Benson et al., 2014).

There was also a significant decrease in BOLD response in regions of the occipital cortex corresponding to retinotopic eccentricities outside those occupied by our stimuli. Such decreases of BOLD response in cortical regions outside of those directly stimulated have been observed previously and have been hypoth-

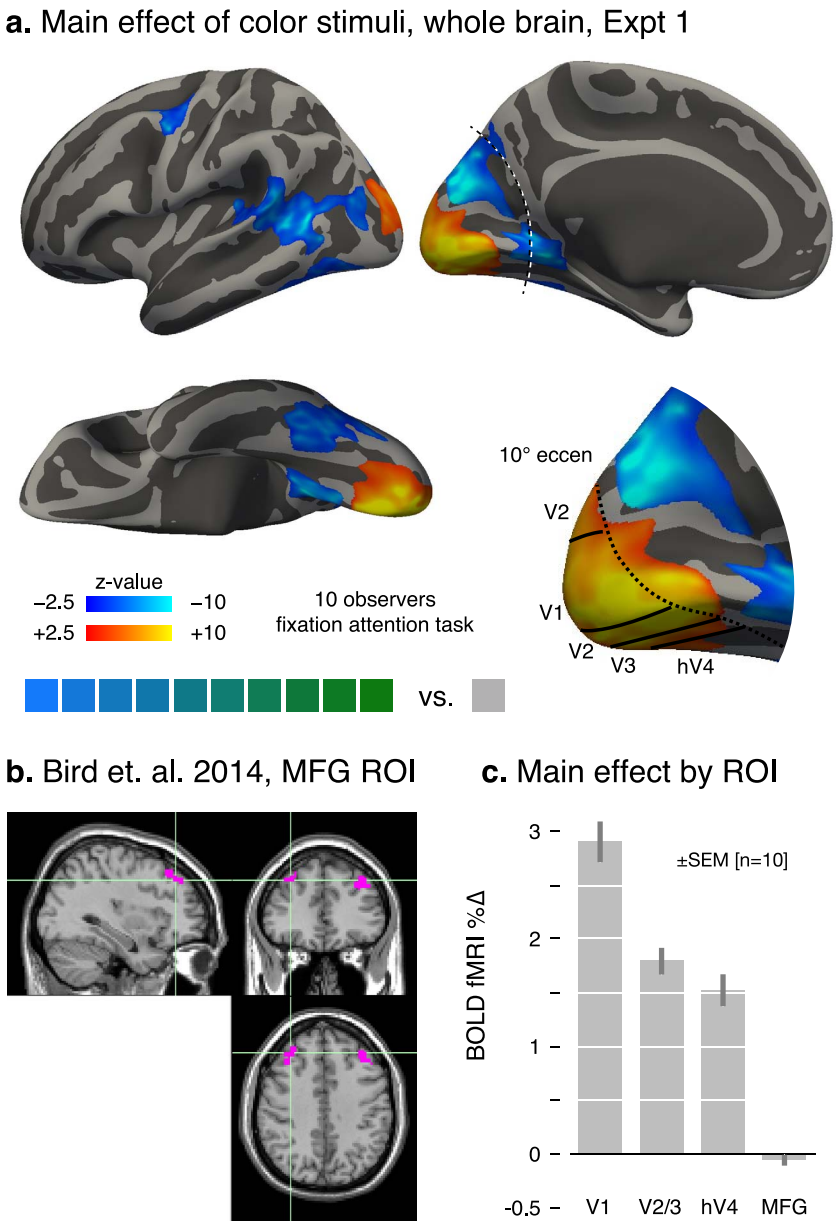

Figure 3. (a) Brain regions where there were significant main effects of all color stimuli versus the uniform background in Experiment 1. These are shown for the aggregate data on the fsaverage-sym, pseudoleft hemisphere surface representation. The inset at the bottom right of the panel shows a magnified view of the occipital cortex with the boundaries between early visual areas marked. The map was thresholded at $|z| > 2.5$, and then permutation was used to set a spatial extent threshold corresponding to a map-wise $p < 0.05$. This same technique was used for all whole-brain maps shown in subsequent figures. (b) Region of interest in the MFG identified by Bird et al. (2014), shown on the standard Montreal Neurological Institute brain. (c) Main effect of color stimuli versus background on BOLD response, shown by ROI. V1, V2/3, and hV4 are visual areas identified using a cortical surface template (Benson et al., 2014). MFG is the ROI identified by Bird et al. (2014).

esized to arise from a surround suppressive mechanism that hyperpolarizes magnocellularly driven neurons at the level of input to V1 (Wade & Rowland, 2010).

Because our main interest is in the representation of color and its categories in early visual cortex, we defined ROIs corresponding to the retinotopically stimulated region (0°–10° eccentricity) of visual areas V1, V2/3, and hV4 and examined in more detail the responses in these areas. The mean BOLD response in each of these ROIs is shown in Figure 3c. The responses in V2 and V3 were combined because of their close physical proximity in the cortex and the absence of an observed difference in the response between these regions.

Bird et al. (2014) concluded that fMRI adaptation reveals a categorical color effect in a region of the MFG,

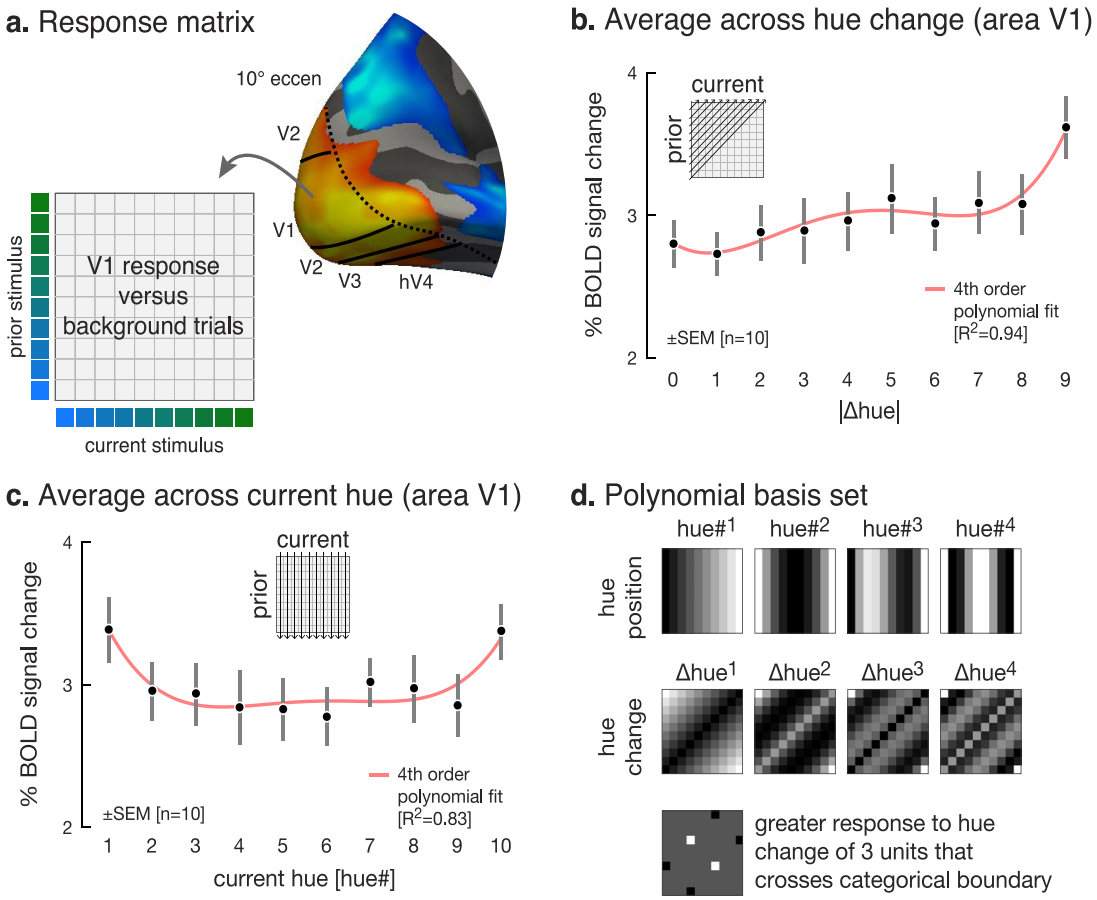

Figure 4. (a) The response matrix. For each ROI, the matrix contains the average BOLD response to each of our color stimuli when it was immediately preceded by each of the other color stimuli. (b) Effect of sequential hue change in V1. Mean percentage change in BOLD response is shown for each size of the absolute value of the hue step between sequential pairs of stimuli, $|\Delta hue|$. The data for each value of $|\Delta hue|$ were aggregated over the possible currently presented stimuli and then over subjects. Error bars show the between-subjects *SEM*. The fit curve is a fourth-order polynomial. This plot does not correct for variation in the direct effect of stimulus hue. (c) Mean BOLD response as a function of the hue of the current stimulus. The data for each current stimulus hue were aggregated over the possible immediately preceding stimulus hues and then over subjects. Error bars show the between-subjects *SEM*. The fit curve is a fourth-order polynomial. This plot does not correct for effects of variation in the size of the hue step between the current and immediately preceding stimulus. (d) Modeling approach using a set of polynomial regressors to simultaneously model the direct effect of stimulus hue and the sequential effect of hue change along with the effect of hue changes of size three that differentially cross the blue–green categorical boundary established by our psychophysical measurements.

but they observed no overall increase in BOLD fMRI signal to their colored stimuli in this region (or in early visual areas). We identified for each of our subjects the voxels corresponding to the region identified by Bird et al. (location kindly provided to us by the authors; Figure 3b). This region was then interrogated in volumetric space in each of our subjects. Consistent with the report of Bird et al. (2014), we found that the mean BOLD response in this region to our colored stimuli is not different from zero (Figure 3c).

Basic features of the adaptation data

Figure 4a illustrates the form of the data we analyzed from each identified region of interest. Each column of the response matrix (Aguirre, 2007) corresponds to one of our 10 color stimuli, arranged in order of nominal hue. Each entry of the column specifies the BOLD response to that colored stimulus when it was immediately preceded by one of the other stimuli. Thus, the lower left entry of the matrix specifies the response to our bluest stimulus when it was immediately preceded by itself, while the upper left entry provides the response to our bluest stimulus when it was immediately preceded by our greenest stimulus. Thus, the response matrix provides the information on how strongly each color stimulus affected the response to each other color stimulus.

In general, we expect fMRI adaptation to result in smaller responses along the main positive diagonal of

ROI	Covariates of no interest only	10-covariate polynomial model	100-covariate response matrix
V1	0.49	0.75	0.77
V2/V3	0.56	0.76	0.78
hV4	0.47	0.58	0.61
MFG	0.72	0.73	0.75

Table 2. Modeled variance for Experiment 1. The average (across subjects) of the proportion of variance explained in the fMRI signal in each region is examined for different models.

the matrix, which correspond to stimuli preceded by themselves, relative to entries far from the diagonal, which correspond to stimuli preceded by stimuli that differ considerably in hue. Intermediate changes in hue may be expected to produce intermediate neural responses, as has been observed in other stimulus domains (e.g., Drucker & Aguirre, 2009). To examine this, we aggregated the responses along each of the positive diagonals of the response matrix. These correspond to the responses for varying values of the difference in hue between two sequentially presented stimuli and are plotted in Figure 4b. As expected, the responses increase quite regularly as a function of $|\Delta\text{hue}|$ (the absolute value of the difference in hue). The data are well fit by a fourth-order polynomial, as shown by the red curve.

Separately, we consider the direct effect on neural response as a function of the current hue presented on a trial. It may be the case that the hues from the ends of the stimulus continuum evoke a different neural response compared with stimuli drawn from the midpoint. The tendency for extreme stimuli to evoke a larger neural response has been reported previously for other stimulus categories, including faces (Freeman, Rule, Adams, & Ambady, 2010; Loffler, Yourganov, Wilkinson, & Wilson, 2005), face silhouettes (Davidenko, Remus, & Grill-Spector, 2012), and abstract shapes (Panis, Wagemans, & Op de Beeck, 2011). This has been referred to as a norm-based effect. Figure 4c shows how the direct response varies with the hue of the stimulus. Here the plotted points were obtained by averaging the columns of the response matrix. The response is elevated for the most extreme blue and green stimuli and is relatively flat across the others. Again, a fourth-order polynomial provides a good approximation to the smooth response across hue position seen in the data.

Although the forms of the responses presented in Figure 4b and c appear conceptually distinct, they are in fact confounded (Aguirre, 2007; Kahn & Aguirre, 2012). This is because not all stimuli enter into the data for each size of hue step. To see this, note that only the bluest (stimulus 1) and greenest (stimulus 10) stimuli contribute to the data for $|\Delta\text{hue}| = 9$, whereas all 10 stimuli contribute to the case $|\Delta\text{hue}| = 1$. Conversely, hues positioned at the ends of the stimulus range tend to participate in larger transitions between stimuli compared with hues in the middle. Thus, the interpretation of the plots shown in Figure 4b and c depends on the relation between the direct and sequential responses to the stimuli. This confounding is an unavoidable property of stimuli arranged along a single perceptual dimension (Kahn & Aguirre, 2012).

Modeling direct and sequential hue change effects

The direct effect of stimulus hue and the sequential effect of hue change may be separated from each other using a regression-based approach (in the context of a sufficiently counterbalanced stimulus sequence; Aguirre, 2007; Kahn & Aguirre, 2012). We took advantage of the smoothness of the data shown in Figure 4b and c and modeled the entire response matrix with nine covariates (Figure 4d) plus a main effect covariate (not pictured). The first four covariates provide a fourth-order polynomial description of the direct hue effect; the second four covariates provide a fourth-order polynomial description of the sequential hue effect.

The final covariate was designed to capture a potential categorical differential in the data. This covariate modeled the differential effect of hue changes of step size three that either did or did not cross the perceptual blue–green category boundary established by our psychophysical measurements (Figure 2a).

The set of 10 covariates (the nine illustrated in Figure 4d plus the main effect covariate) was fit to the data from each subject and ROI. Below we refer to the full set of 10 covariates as the 10-covariate model. To the extent that the covariates describe each fit response matrix, the fitting allowed us to separate direct and sequential effects and avoid the confounds present in the averaged data shown in Figure 4b and c.

We assessed the performance of the 10-covariate model within each region compared with a 100-covariate model that accounted for all possible elements of the response matrix (Table 2). The proportion of the variance of the fMRI signal in each region explained by the common nuisance covariates of no interest (e.g., global signal effects, head motion) was approximately 0.5 in the visual areas and 0.72 in the MFG. The 10-covariate model and the 100-covariate model each explained an additional 0.15 to 0.25 of variance in the visual areas and 0.02 to 0.04 within the MFG (Table 2). The minimal difference in explanatory power between the models was on the order of 0.02 of the variance. We concluded that the 10-covariate model provides an adequate account of the neuroimaging data and that there are not large, unmodeled, systematic effects of the stimuli in the data.

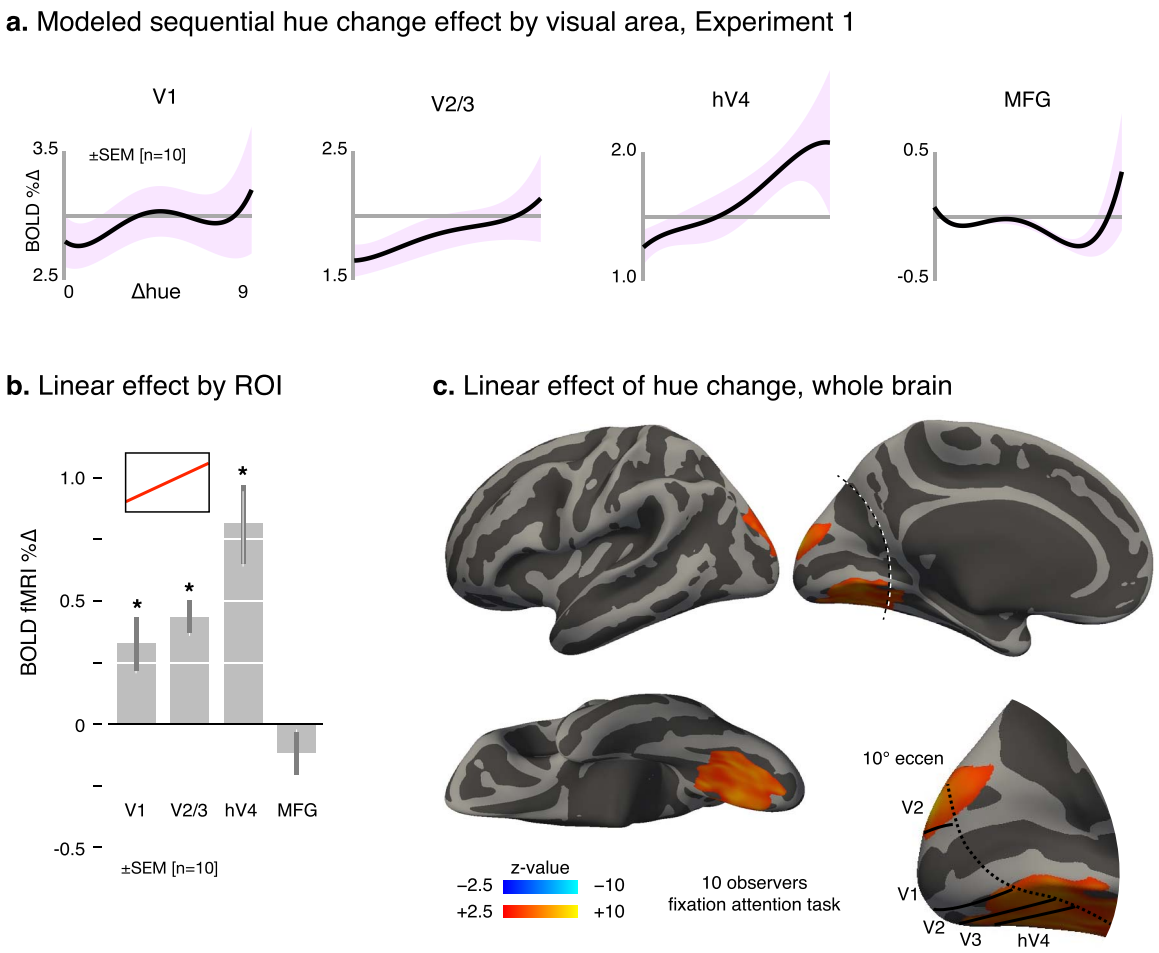

Figure 5. Sequential hue effects in Experiment 1. (a) Modeled sequential hue change effect in V1, V2/3, hV4, and MFG. The data for V2 and V3 are combined because they showed very similar responses. Each curve is obtained from the regression weights on the four sequential effect covariates (Figure 4d) and is thus a fourth-order polynomial. (b) The linear sequential hue effect by ROI. V1: $t(9\ df) = 2.69$, $p = 0.02$; V2/V3: $t(9\ df) = 5.6$, $p = 0.0003$; hV4: $t(9\ df) = 4.22$, $p = 0.002$; MFG: $t(9\ df) = -0.58$, $p = 0.58$. (c) Significance map of the linear effect of sequential hue change on BOLD fMRI response across the whole brain.

BOLD fMRI response to sequential hue change

Figure 5a shows the modeled sequential hue effect by ROI. Each plot in this panel is in the same format as the plot shown in Figure 4b, but here the confound with direct effects has been removed. The overall pattern is consistent with that shown in Figure 4b: The suppression of BOLD fMRI signal decreases with increasing hue distance between successive stimuli. This effect is seen in V1, V2/V3, and hV4. In the MFG ROI, however, there is no such sequential hue effect, although there does appear to be a release from adaptation at the largest hue step sizes. The lack of a linear effect in MFG is perhaps not surprising, given that there was no main effect of the stimuli within this ROI, and is consistent with the reported findings of Bird et al. (2014).

Figure 5b shows the loading on the sequential hue linear covariate by ROI. The linear covariate was significant in all visual areas. In a post hoc test, the loading on the linear covariate across subjects was found to be greater in hV4 compared with V1, $t(9\ df) = 3.08$, $p = 0.01$. There was no significant loading on any of the higher polynomial components of the response, with the exception of a quartic (fourth-order polynomial) effect in area V1 ($p = 0.02$). There was no significant response to sequential hue change for any of the modeled components within the MFG ROI. Figure 5c shows a whole-brain map of the loading on the linear covariate. There is a consistent sequential effect of adaptation that decreases linearly with the increasing hue change (with respect to our hue scale) in early visual cortex and no apparent effect of adaptation elsewhere in the brain. While there was a significant linear response within the V1 ROI, the spatial extent of this response was too small to exceed the spatial extent threshold used to generate the whole-brain map. No significant responses were observed at the whole-brain level for any of the higher polynomial components that modeled neural response to sequential hue change. The

linear effect we observe suggests that the neural representation of color in the corresponding visual areas is sensitive to continuous variation in the color difference between sequentially presented stimuli during passive viewing of the color stimuli.

BOLD fMRI response to categorical hue change

While the results presented in Figure 5 confirm the sensitivity of early visual areas to hue, the measures considered combine data across sequential hue changes that do and do not cross the blue–green category boundary. The signature of a categorical representation would be a differential recovery from adaptation for stimulus changes that cross the categorical boundary versus those that do not. This type of effect has previously been observed in extrastriate occipital areas for a learned object category (Folstein, Palmeri, & Gauthier, 2013) and is what Bird et al. (2014) examined in their study.

To assess whether there was a categorical sequential hue effect in addition to the linear sequential hue effect, we included a covariate that modeled the categorical effect of a change in hue of three units that either did or did not cross the perceptual categorical boundary and examined the loading on this covariate. The covariate contrasted the BOLD signal measured on trials for sequential hue steps of size three that were within and between the blue–green color categories, as indicated at the top of Figure 6. We used a step size of three units because this is the only transition size that was large enough to reliably cross the categorical boundary in all subjects for one pair of stimuli while still small enough to remain within category for other pairs (see Figure 2). The same covariate was used for all subjects.

Figure 6 shows the results for Experiment 1 by ROI. In early visual areas, the loadings on the category covariate are all negative (in the direction opposite to that predicted by a category effect) and none are statistically significant. In the MFG, there is a nonsignificant positive loading. A whole-brain analysis revealed no significant loading anywhere in the brain. Our data thus provide no evidence in support of a categorical representation in the early visual cortex or in the MFG ROI of Bird et al. (2014) when subjects performed a fixation-attention task. In the second experiment, we considered the possibility that this null result occurred because subjects were not actively providing category labels for the stimuli.

BOLD fMRI response to hue position

In addition to measuring a linear release from adaptation in early visual areas, we measured the direct effect of each color stimulus during passive viewing. Again, we modeled the data from the passive-viewing

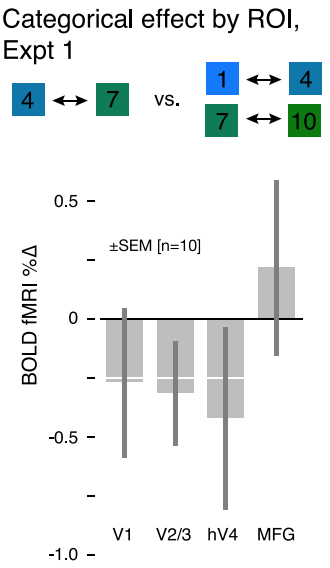

Figure 6. Categorical adaptation effects in Experiment 1. The top panel illustrates the steps that were contrasted by our categorical effect covariate. The bottom panel shows the loading on this categorical covariate by ROI. None of the loadings reach statistical significance. V1: $t(9\ df) = -0.82$, $p = 0.33$; V2/V3: $t(9\ df) = -1.34$, $p = 0.21$; hV4: $t(9\ df) = -0.99$, $p = 0.34$; MFG: $t(9\ df) = 1.01$, $p = 0.34$.

scan within our retinotopically defined visual areas and MFG, using the polynomial basis set described above. There was a greater response to the extreme stimuli in V1, V2/3, and hV4 (Figure 7a). This is revealed by statistically significant quadratic and quartic (fourth-order polynomial) effects in these visual areas (Figure 7b, c). An interesting property of these data is that there appears to be an enhanced response to the hues in the middle of the stimulus range that develops between visual areas V1 and hV4. This is reflected in the declining loading on the quadratic covariate and rising loading on the quartic covariate across visual areas. A post hoc test that examined the difference in loading between the quartic and quadratic covariates across subjects and between areas hV4 compared with V1 was $t(9\ df) = 2.10$, $p = 0.07$.

Subjects were not engaged in a categorization task during the collection of the neuroimaging data (and indeed were instead attending to changes in an achromatic fixation dot). Nonetheless, it is interesting to note that the rising response to midpoint hues in area hV4 is similar to the longer response times needed to categorize midpoint stimuli as blue or green in the separate behavioral data (Figure 2c). The enhanced direct response to midpoint stimuli in hV4 may represent the activity of a categorization process developing at this cortical level, which requires more neural processing to assign a category to ambiguous stimuli.

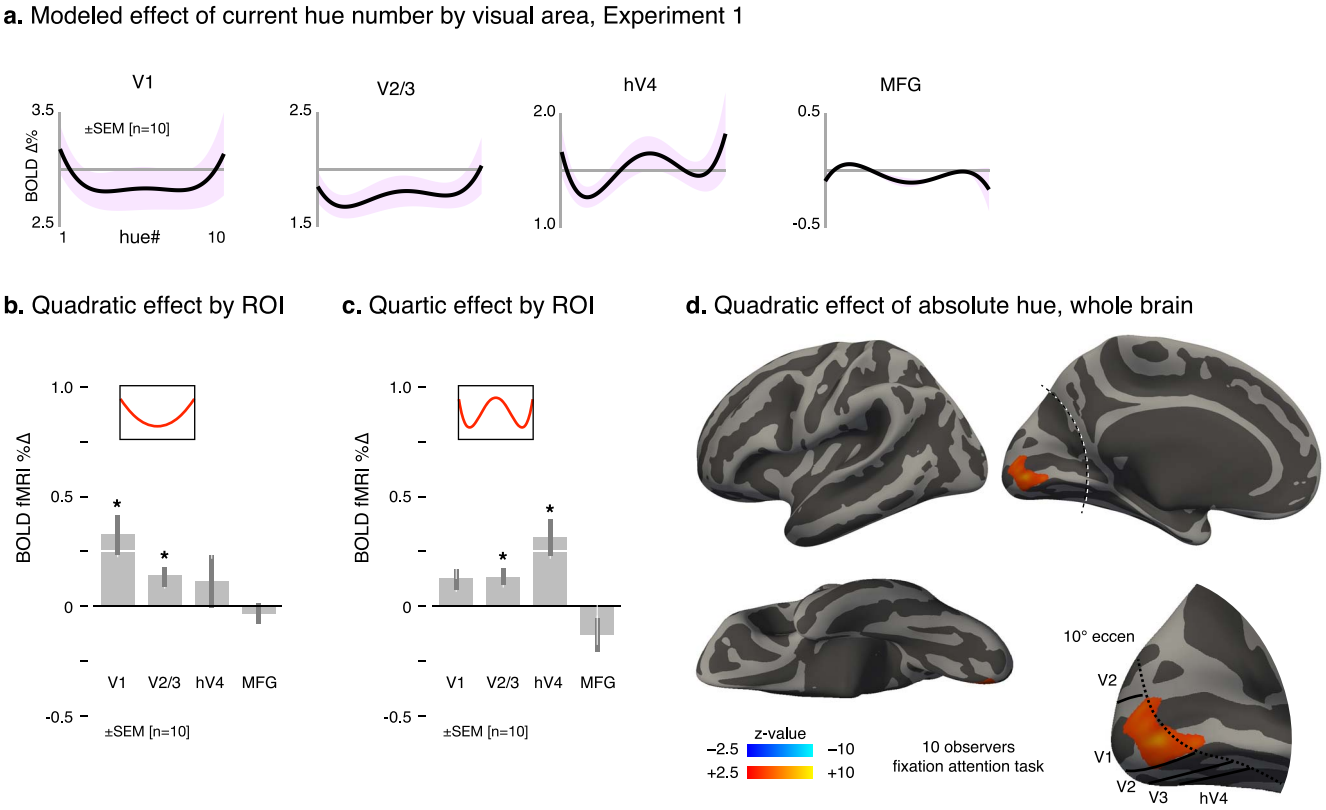

Figure 7. Direct hue effects in Experiment 1. (a) Modeled direct effect of hue in our ROIs. (b) Quadratic effect by ROI. V1: $t(9\ df) = 3.17$, $p = 0.01$; V2/V3: $t(9\ df) = 2.42$, $p = 0.04$; hV4: $t(9\ df) = 0.81$, $p = 0.44$; MFG: $t(9\ df) = -0.34$, $p = 0.74$. (c) Quartic effect by ROI. V1: $t(9\ df) = 2.05$, $p = 0.07$; V2/V3: $t(9\ df) = -2.71$, $p = 0.02$; hV4: $t(9\ df) = 2.99$, $p = 0.01$; MFG: $t(9\ df) = -1.54$, $p = 0.16$. (d) Whole-brain map of significant quadratic effects.

The voxels studied here were identified as those having a significant main effect response to the entire set of hues compared with the background trials. We note that it is possible in principle that this thresholding by the main effect could alter the form of the reported direct effect responses, biasing against finding responses in which there is a small or even negative response to some hues. In practice, the large and confluent main effect responses seen within the studied regions of interest (Figure 3a) minimize this concern.

Experiment 2

Method

We considered that our inability to find categorical representations of color in Experiment 1 was due to the nature of the fixation-monitoring task used during the scanning. It could be that categorical representations in the early visual cortex arise only when subjects actively categorize colors. During active color naming, do categorical representations emerge in early visual areas? Furthermore, can we replicate with active viewing the categorical adaptation effect in MFG reported by Bird et al. (2014) for passive viewing? To test this hypothesis, we repeated our experiment with a subset of five subjects, but during Experiment 2 participants actively categorized the color stimuli during scanning.

Five subjects from Experiment 1 returned to participate in Experiment 2, which was conducted 9 to 10 months after Experiment 1. Scanning parameters and fMRI analyses were identical to those in Experiment 1. Here, however, subjects actively categorized colors by indicating via a left or right button press for each checkerboard stimulus whether the color was blue or green. The response mapping was randomized across subjects (i.e., for some subjects the right button indicated *blue* and for others it indicated *green*). Responses for each trial were accepted if made within the 1100-ms window that began with stimulus onset and ended with the onset of the subsequent stimulus.

Results

Participants reliably categorized the colors as either blue or green during scanning (Supplementary Figure

ROI	Covariates of no interest only	10-covariate polynomial model	100-covariate response matrix
V1	0.45	0.74	0.75
V2/V3	0.53	0.77	0.79
hV4	0.54	0.68	0.70
MFG	0.74	0.78	0.79

Table 3. Modeled variance for Experiment 2. The average (across subjects) of the proportion of variance explained in the fMRI signal in each region is examined for different models.

S1) but with more lapses than in Experiment 1. This may be because the timing of the response was governed by the requirements of the scan itself rather than being paced by the subject. In addition to the lapses, subjects failed to respond to a substantial number of stimuli by the response deadline for each trial. Overall, however, their categorization boundaries were similar to what we found in Experiment 1.

We analyzed the neuroimaging data in the same fashion as for Experiment 1. We again confirmed that the polynomial model fit the data well (Table 3). Figure 8a shows the sequential effects of hue change by ROI, obtained via our 10-covariate regression approach. There are some similarities to the responses found in Experiment 1. A significant linear effect is found in early visual areas and not in the MFG (Figure 8b). These linear effects are found at the whole-brain level only within the visual cortex. Unlike Experiment 1, however, the active-categorization task is associated with a flattening of the recovery from adaptation at higher stimulus change levels. This is evidenced as a negative loading on the quadratic component of the response (Figure 8c) in early visual areas. A direct comparison of the data for subjects in the two experiments (paired, two-tailed t test) showed a significant difference in the loading on the quadratic covariate only in the V2/V3 ROI; V1: $t(4\ df) = 0.16$, $p = 0.88$; V2/V3: $t(4\ df) = 2.8$, $p = 0.05$; hV4: $t(4\ df) = 1.69$, $p = 0.17$.

We also examined the direct effect of hue position on the BOLD fMRI signal (Figure 9). Notable is the

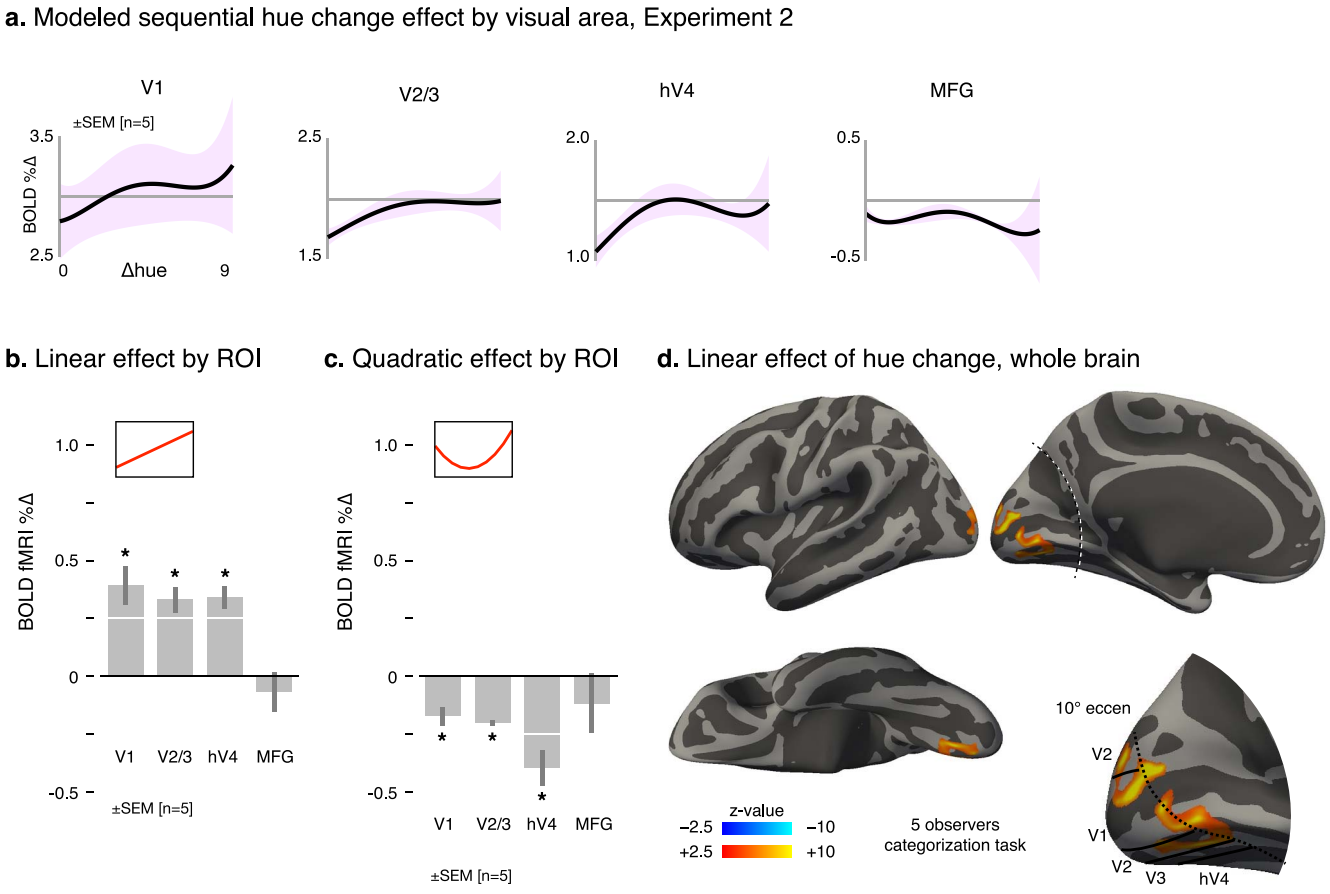

Figure 8. Sequential hue effects in Experiment 2. (a) Modeled sequential hue change effect in V1, V2/3, hV4, and MFG (same format as Figure 5a). (b) The linear sequential hue effect by ROI. V1: $t(9\ df) = 3.00$, $p = 0.02$; V2/V3: $t(9\ df) = 3.37$, $p = 0.008$; hV4: $t(9\ df) = 3.64$, $p = 0.005$; MFG: $t(9\ df) = -0.63$, $p = 0.55$. (c) The quadratic sequential hue effect by ROI. V1: $t(9\ df) = 2.51$, $p = 0.03$; V2/V3: $t(9\ df) = 4.62$, $p = 0.001$; hV4: $t(9\ df) = 3.16$, $p = 0.01$; MFG: $t(9\ df) = -0.80$, $p = 0.44$. (d) Significance map of the linear effect of sequential hue change on BOLD fMRI response across the whole brain.

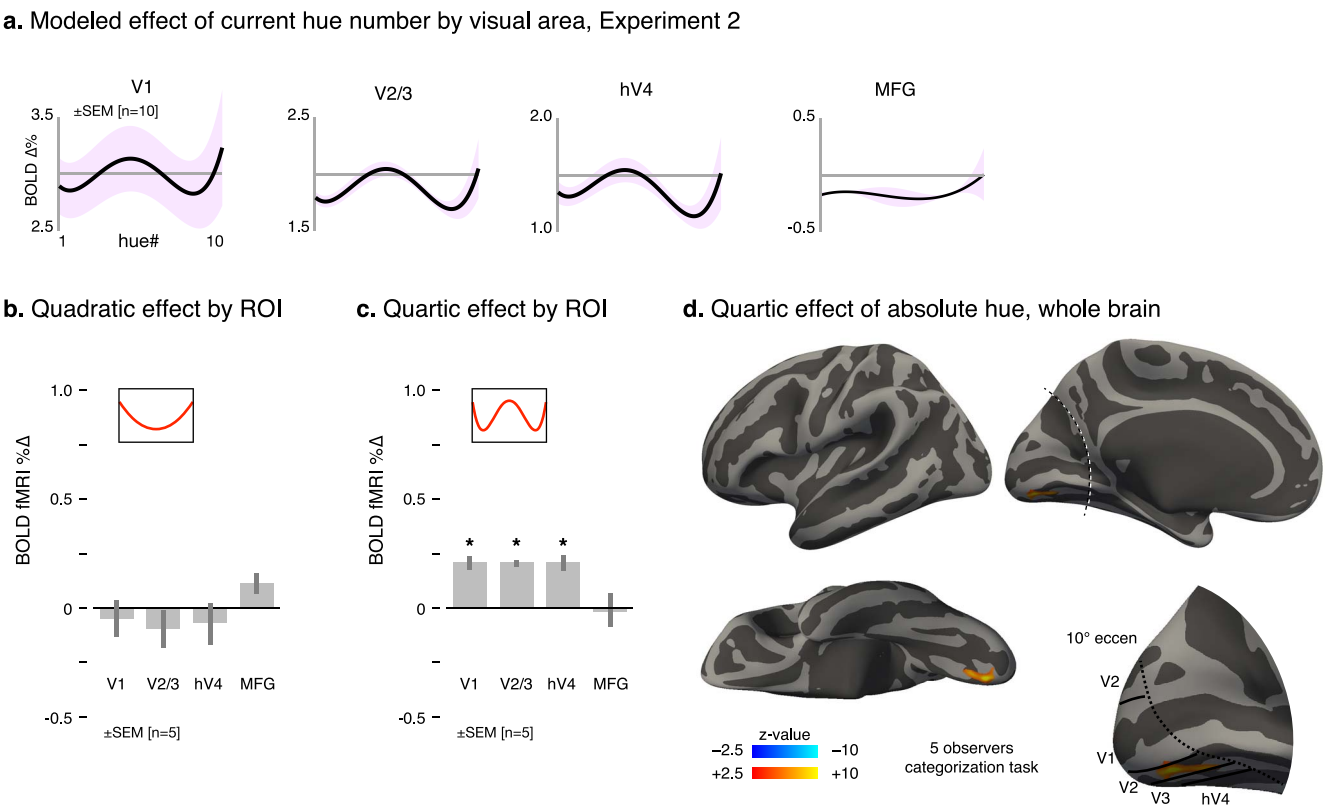

Figure 9. Direct hue effects in Experiment 2. (a) Modeled direct effect of hue in our ROIs. (b) Quadratic effect by ROI. V1: $t(9\ df) = -0.44$, $p = 0.67$; V2/V3: $t(9\ df) = -0.88$, $p = 0.44$; hV4: $t(9\ df) = -0.57$, $p = 0.58$; MFG: $t(9\ df) = 1.70$, $p = 0.12$. (c) Quartic effect by ROI. V1: $t(9\ df) = -3.35$, $p = 0.01$; V2/V3: $t(9\ df) = 4.29$, $p = 0.002$; hV4: $t(9\ df) = 3.00$, $p = 0.01$; MFG: $t(9\ df) = -0.04$, $p = 0.97$. (d) Whole-brain map of significant quartic effects.

general enhancement of loading on the quartic (fourth-order polynomial) component. As with the hV4 data from Experiment 1, this is revealed as an enhanced response to stimuli from the midpoint of the stimulus range, now present in all early visual areas. A paired t test did not find this, however, to be a significant change from the quartic response in Experiment 1 for any of the early visual areas; V1: $t(4\ df) = 0.39$, $p = 0.72$; V2/V3: $t(4\ df) = 0.31$, $p = 0.77$; hV4: $t(4\ df) = 0.93$, $p = 0.41$.

Finally, we used our covariate that modeled step sizes of three that either did or did not cross the categorical boundary to model categorical adaptation effects across the whole brain. We did not find a categorical effect in early visual areas or in the MFG ROI provided by Bird et al. (2014; Figure 10).

We considered whether the lack of a categorical effect in early visual areas occurred because the study was insufficiently powered. This seems unlikely, however, as the main effect of stimuli (Supplementary Figure S2) was similar to that observed in Experiment 1 (Figure 3).

Next, we considered whether the lack of a categorical effect within the MFG was due to imperfect cross-study brain alignment so that we might have missed a small, localized effect close to the ROI provided to us by the authors of Bird et al. (2014). We conducted an exploratory whole-brain analysis using our categorical covariate. As noted above, this analysis revealed no brain regions with significant loading in Experiment 1. In Experiment 2, however, three frontal brain regions did reveal a significant categorical covariate loading (Figure 10b). These were identified in an analysis in which the data from the two hemispheres were combined in a single, pseudoleft hemisphere. No significant areas of response were identified by this analysis when the data from the two hemispheres were not combined. These were spatially separated in the frontal lobe, appearing in the precentral sulcus, the pars triangularis of the inferior frontal gyrus, and the posterior part of the superior frontal sulcus. The bilateral region identified Bird et al. (2014) was substantially more anterior or superior than any of the areas identified in our study. To conclude, we did not find a categorical adaptation effect to color stimuli anywhere in the brain during passive viewing. During active color naming, we found no categorical effect in early visual areas. An exploratory analysis, however, revealed a category effect in several frontal areas, albeit

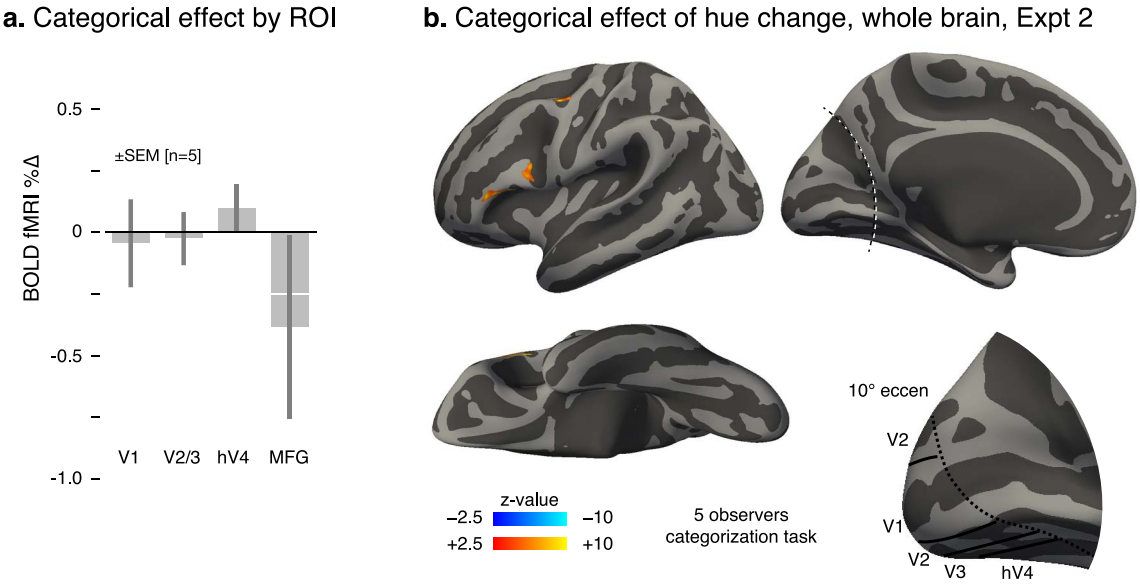

Figure 10. Categorical hue adaptation effects in Experiment 2. (a) Categorical effect by ROI. V1: $t(9\ df) = -0.21$, $p = 0.84$; V2/V3: $t(9\ df) = -0.17$, $p = 0.87$; hV4: $t(9\ df) = 0.84$, $p = 0.42$; MFG: $t(9\ df) = -0.93$, $p = 0.38$. (b) Whole-brain map of regions showing significant loading on the three-step categorical covariate.

ones that are at some distance from the site reported by Bird et al. (2014).

Discussion

We used an event-related, continuous carryover fMRI adaptation design to interrogate the representation of color in early visual cortex during both passive viewing (Experiment 1) and active categorization (Experiment 2) of color stimuli. We defined an additional ROI in the MFG, based on a report of categorical encoding of color in this region (Bird et al., 2014). We simultaneously modeled the direct effect of stimulus hue and the sequential effect of hue change along with the effect of hue changes that differentially cross the blue–green categorical boundary established by our psychophysical measurements. We emphasize that such simultaneous modeling is necessary to disentangle the separate contributions of direct, sequential, and categorical effects on the data represented by the measured response matrix.

Our results from both experiments demonstrate in visual cortex a smoothly increasing recovery from adaptation with increasing hue distance between adjacent stimuli. This effect is present within area V1 but grows in amplitude and reliability through V2/V3 and into area hV4. In whole-brain maps, the effect is notably localized to early extrastriate visual cortex, extending slightly beyond area hV4 into the ventral occipital cortex. Prior studies have used fMRI adaptation to demonstrate the presence of color representation within early visual cortex (Engel, 2005; Engel & Furmanski, 2001). To our knowledge, our measurements represent the first demonstration that fMRI adaptation may be used to probe the parametric form of the neural representation of color.

Our stimuli varied along the dimension of Munsell hue. This choice was motivated by the fact that subjects can readily use color terms to categorize stimuli that vary in this manner. We focused on stimuli in the green–blue hue range but could have chosen other regions of the hue circle that are also easily divided into two color categories (e.g., blue–purple, orange–yellow). We have no reason to believe that the qualitative features of our results would fail to generalize to other regions of the hue circle, but we have not investigated this question. In addition, we note that varying hue is not the only way to parametrically change the color appearance of stimuli. It would certainly be possible to construct alternative stimulus sets (e.g., one in which Munsell chroma or value was varied and hue was held constant). For such stimuli, the change in appearance is unlikely to be well described using basic color terms. There is no guarantee that the qualitative patterns of fMRI BOLD direct and adaptation effects we observed would be seen for such stimulus sets. For example, varying Munsell chroma might well have a large direct effect on the measured response, as this will change the salience of the stimuli relative to the background.

The color-tuning properties of neurons in the early visual cortex have been well studied (for reviews see Conway, 2014; Gegenfurtner, 2003; Shapley & Hawken, 2011; Solomon & Lennie, 2007). Relating this literature to our results is complicated by the hetero-

geneity of findings across laboratories (Shapley & Hawken, 2011), by the fact that single units combine signals from cones nonlinearly (Horwitz & Hass, 2012; Solomon, Peirce, & Lennie, 2004), and by the fact that the response of individual neurons depends on the spatial and temporal structure of the stimulus (Shapley & Hawken, 2011; Solomon & Lennie, 2007). Broadly speaking, at least some individual neurons in early visual cortex tend to respond selectively to a small range of hues when other stimulus features are held fixed, making plausible the notion that hue could be represented by a population code (Wachtler, Sejnowski, & Albrecht, 2003). Our results are consistent with such a representation: As the hue distance between stimuli increases, the number of neurons driven strongly by the two stimuli would be expected to decrease, and thus we would expect a falloff in the adaptation effect of the sort we observe. Our stimulus set, however, was not designed to probe the dependence of fMRI adaptation on variation in all three dimensions of color space, and thus our data do not place strong constraints on the nature of the underlying neural code.

Our measurements were made using a particular timing of presentation of stimuli. Cortical visual areas may differ in their sensitivity to adaptation at different time scales. Therefore, more rapid or slow presentations of our stimuli might produce a different pattern of recovery from adaptation across visual areas. In future studies, varying the timing of stimulus presentation might be used to examine the differential sensitivity of cortex to the temporal integration of visual information. These measurements may also be linked to perceptual manipulations of stimulus discriminability and similarity.

We did not find in visual cortex a difference in recovery of adaptation for hue changes that did or did not cross the blue–green categorical boundary. This was true during both passive viewing and active categorization. Such an effect has been observed in the extrastriate cortex for stimuli consisting of shapes that spanned a learned categorical boundary (Folstein et al., 2013). Our results support the view that the categorical representation of color emerges in areas beyond hV4. We also note that, had we obtained a significant categorical effect, additional controls would have been necessary to rule out the possibility that the cause of the effect was inadvertent irregular perceptual spacing of the stimuli. Although our stimuli were chosen from a nominally perceptually uniform space, such spaces are well known to only approximate their design goal of perceptual uniformity (Brainard, 2003).

Bird et al. (2014) reported a categorical fMRI adaptation effect in the bilateral MFG under conditions in which subjects performed an attention task not related to the color of the stimuli. In contrast, we did not find a categorical representation of color in this brain region in either Experiment 1 or Experiment 2. We considered the possibility that we missed the categorical effect that they reported because of imperfect alignment of brain regions across studies. In an exploratory whole-brain analysis of the data from Experiment 2, we did find significant loading on the categorical covariate in three frontal clusters. While this would seem to support the broad conclusion drawn by Bird et al. (2014), we regard both our finding and that of Bird et al. as quite tentative because the regions in both studies were discovered using a whole-brain analysis conducted with a relatively small number of subjects, for which cluster control of the map-wise false positive rate is notoriously suspect (Woo, Krishnan, & Wager, 2014).

An important difference between our study and that of Bird et al. (2014) is the sequencing of presented stimuli. In our study, subjects viewed a continuous stream of counterbalanced stimuli that had minimal discernable temporal structure. The design of Bird et al. had subjects view approximately 10-s blocks of pairs of alternating color stimuli. Across blocks, the pairs either differed or matched in categorical color name assignment. A feature of this design is that categorical transitions were explicitly grouped in the stimulus sequence. Thus, for example, subject awareness of this grouping could be responsible for an MFG response, in keeping with other studies that have generally found responses in this area to explicit categorical manipulations (e.g., Myers & Swan, 2012). As Bird et al. did not find a main effect response to the color stimuli in their frontal region (replicated in our study), we suspect that any categorical response in this region (if replicated) is domain general and not specific to color representation.

It is possible that a categorical representation of color labels is present in portions of the temporal lobes that were imperfectly imaged in this study. Patients with semantic dementia (a progressive neurodegenerative disorder associated with atrophy of the anterior temporal lobe) have general impairments in the representation of categories of semantic information. Within the domain of color, these patients exhibit impairments in color naming despite preserved ability to discriminate color (Rogers, Graham, & Patterson, 2015; Rogers, Patterson, & Graham, 2007). As the anterior temporal lobes are poorly imaged using echoplanar fMRI due to susceptibility artifacts, it is possible that a categorical response to our stimuli was present but unmeasured in this brain region.

In contrast to the absence of categorical adaptation effects, we did observe an increased direct BOLD fMRI response to stimuli from the midpoint of the hue range, near the categorical boundary. This was observed in hV4 in Experiment 1 and in all visual areas in Experiment 2. The form of this effect is reminiscent of

the form of the dependence of response time on hue (compare Figures 7a and 9a with Figure 2c). The enhanced direct response to midpoint stimuli in hV4 may be indicative of a categorization process developing at this cortical level. Brouwer and Heeger (2013) used multivoxel pattern analysis to derive the structure of the neural representation of color. In hV4 and VO1, they observed an increased categorical clustering when subjects actively categorized the stimuli. They offered a model of this effect that posited differential task-dependent gains for stimuli at the center and edges of color categories—an idea that has some support from single unit studies in awake, behaving monkeys (Koida & Komatsu, 2007). Our data are consistent with such a differential gain change, but in our case the increased gain is for the stimuli near the category boundary.

In summary, there is a smooth, parametric relationship between the perceptual dissimilarity of hue pairs and the magnitude of fMRI signal evoked by the transitions between them. The neural similarities that we measured using fMRI sequential adaptation suggest that the representation of color in the visual cortex is fine grained and that color categories do not intrude on this representation. The change in the dependence of direct effect on stimulus hue, however, leaves open the possibility that response gain changes that accompany a color-categorization task may be manifest in these same visual areas.

Keywords: color, categorization, functional magnetic resonance imaging

Acknowledgments

This research was supported by National Institutes of Health Grants EY10016 (D. H. B.) and EY021717 (S. T. S. and G. K. A.). Christopher Broussard provided technical assistance. We thank Anna Franklin and Chris Bird for comments on the article.

Commercial relationships: none.
Corresponding authors: David H. Brainard; Geoffrey K. Aguirre.
Email: brainard@psych.upenn.edu; aguirreg@mail.med.upenn.edu.
Address: Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA; Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA.

References

Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *NeuroImage, 35,* 1480–1494, doi:10.1016/j.neuroimage.2007.02.005.

Aguirre, G. K., Mattar, M. G., & Magis-Weinberg, L. (2011). de Bruijn cycles for neural decoding. *NeuroImage, 56,* 1293–1300, doi:10.1016/j.neuroimage.2011.02.005.

Aguirre, G. K., Zarahn, E., & D'esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage, 8,* 360–369, doi:10.1006/nimg.1998.0369.

Benson, N. C., Butt, O. H., Brainard, D. H., & Aguirre, G. K. (2014). Correction of distortion in flattened representations of the cortical surface allows prediction of V1-V3 functional organization from anatomy. *PLoS Computational Biology, 10*(3), e1003538, doi:10.1371/journal.pcbi.1003538.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.

Bird, C. M., Berens, S. C., Horner, A. J., & Franklin, A. (2014). Categorical encoding of color in the brain. *Proceedings of the National Academy of Sciences, USA, 111,* 4590–4595, doi:10.1073/pnas.1315275111.

Brainard, D. H. (2003). Color appearance and color difference specification. In S. K. Shevell (Ed.), *The science of color, 2nd edition* (pp. 191–216). Washington, DC: Optical Society of America.

Brainard, D. H., Pelli, D. G., & Robson, T. (2002). Display characterization. In J. P. Hornak (Ed.), *Encyclopedia of imaging science and technology* (pp. 172–188). New York: Wiley.

Brouwer, G. J., & Heeger, D. J. (2013). Categorical clustering of the neural representation of color. *Journal of Neuroscience, 33,* 15454–15465, doi:10.1523/JNEUROSCI.2472-13.2013.

Clifford, A., Franklin, A., Holmes, A., Drivonikou, V. G., Özgen, E., & Davies, I. R. L. (2012). Neural correlates of acquired color category effects. *Brain and Cognition, 80,* 126–143, doi:10.1016/j.bandc.2012.04.011.

Conway, B. R. (2014). Color signals through dorsal and ventral visual pathways [Special issue]. *Visual Neuroscience, 31,* 197–209.

Cross, D. V., Lane, H. L., & Sheppard, W. C. (1965). Identification and discrimination functions for a visual continuum and their relation to the motor theory of speech perception. *Journal of Experimental Psychology, 70,* 63–74, doi:10.1037/h0021984.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical

surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage, 9,* 179–194.

Davidenko, N., Remus, D. A., & Grill-Spector, K. (2012). Face-likeness and image variability drive responses in human face-selective ventral regions. *Human Brain Mapping, 33,* 2334–2349, doi:10.1002/hbm.21367.

Drucker, D. M., & Aguirre, G. K. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cerebral Cortex, 19,* 2269–2280, doi:10.1093/cercor/bhn244.

Engel, S. A. (2005). Adaptation of oriented and unoriented color-selective neurons in human visual areas. *Neuron, 45,* 613–623, doi:10.1016/j.neuron.2005.01.014.

Engel, S. A., & Furmanski, C. S. (2001). Selective adaptation to color contrast in human primary visual cortex. *Journal of Neuroscience, 21,* 3949–3954.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences, USA, 97,* 11050–11055.

Fischl, B., Sereno, M., & Dale, A. (1999). Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage, 9,* 195–207, doi:10.1006/nimg.1998.0396.

Folstein, J. R., Palmeri, T. J., & Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex, 23,* 814–823, doi:10.1093/cercor/bhs067.

Freeman, J. B., Rule, N. O., Adams, R. B., & Ambady, N. (2010). The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. *Cerebral Cortex, 20,* 1314–1322, doi:10.1093/cercor/bhp195.

Gegenfurtner, K. (2003). Cortical mechanisms of colour vision. *Nature Neuroscience, 4,* 563–572.

Greve, D. N., Van der Haegen, L., Cai, Q., Stufflebeam, S., Sabuncu, M. R., Fischl, B., & Brysbaert, M. (2013). A surface-based analysis of language lateralization and cortical asymmetry. *Journal of Cognitive Neuroscience, 25,* 1477–1492, doi:10.1162/jocn_a_00405.

Hagler, D. J., Saygin, A. P., & Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage, 33,* 1093–1103, doi:10.1016/j.neuroimage.2006.07.036.

He, X., Witzel, C., Forder, L., Clifford, A., & Franklin, A. (2014). Color categories only affect post-perceptual processes when same-and different-category colors are equally discriminable. *Journal of the Optical Society of America A, 31,* A322–A331.

Horwitz, G. D., & Hass, C. A. (2012). Nonlinear analysis of macaque V1 color tuning reveals cardinal directions for cortical color processing. *Nature Neuroscience, 15,* 913–919.

Ishihara, S. (1936). series of plates designed as tests for colour-blindness. Location: Publisher.

Kahn, D. A., & Aguirre, G. K. (2012). Confounding of norm-based and adaptation effects in brain responses. *NeuroImage, 60,* 2294–2299, doi:10.1016/j.neuroimage.2012.02.051.

Koida, K., & Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nature Neuroscience, 10,* 108–116, doi:10.1038/nn1823.

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience, 8,* 1386–1391, doi:10.1038/nn1538.

Myers, E. B., & Swan, K. (2012). Effects of category learning on neural sensitivity to non-native phonetic categories. *Journal of Cognitive Neuroscience, 24,* 1695–1708.

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15*(1), 1–25.

Panis, S., Wagemans, J., & Op de Beeck, H. P. (2011). Dynamic norm-based encoding for unfamiliar shapes in human visual cortex. *Journal of Cognitive Neuroscience, 23,* 1829–1843, doi:10.1162/jocn.2010.21559.

Rogers, T. T., Graham, K. S., & Patterson, K. (2015). Semantic impairment disrupts perception, memory, and naming of secondary but not primary colours. *Neuropsychologia, 70,* 296–308.

Rogers, T. T., Patterson, K., & Graham, K. (2007). Colour knowledge in semantic dementia: It is not all black and white. *Neuropsychologia, 45,* 3285–3298.

Shapley, R., & Hawken, M. J. (2011). Color in the cortex: Single- and double-opponent cells. *Vision Research, 51,* 701–717.

Solomon, S. G., & Lennie, P. (2007). The machinery of colour vision. *Nature Reviews Neuroscience, 8,* 276–286.

Solomon, S. G., Peirce, J. W., & Lennie, P. (2004). The impact of suppressive surrounds on chromatic

properties of cortical neurons. *Journal of Neuroscience, 24,* 148–160.

Wachtler, T., Sejnowski, T. J., & Albrecht, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron, 37,* 681–691.

Wade, A. R., & Rowland, J. (2010). Early suppressive mechanisms and the negative blood oxygenation level-dependent response in human visual cortex. *Journal of Neuroscience, 30,* 5008–5019, doi:10.1523/JNEUROSCI.6260-09.2010.

Witthoft, N., Nguyen, M. L., Golarai, G., LaRocque, K. F., Liberman, A., Smith, M. E., & Grill-Spector, K. (2014). Where is human V4? Predicting the location of hV4 and VO1 from cortical folding. *Cerebral Cortex, 24,* 2401–2408, doi:10.1093/cercor/bht092.

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage, 91,* 412–419, doi:10.1016/j.neuroimage.2013.12.058.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—Again. *NeuroImage, 2,* 173–181, doi:10.1006/nimg.1995.1023.

Supplemental figures

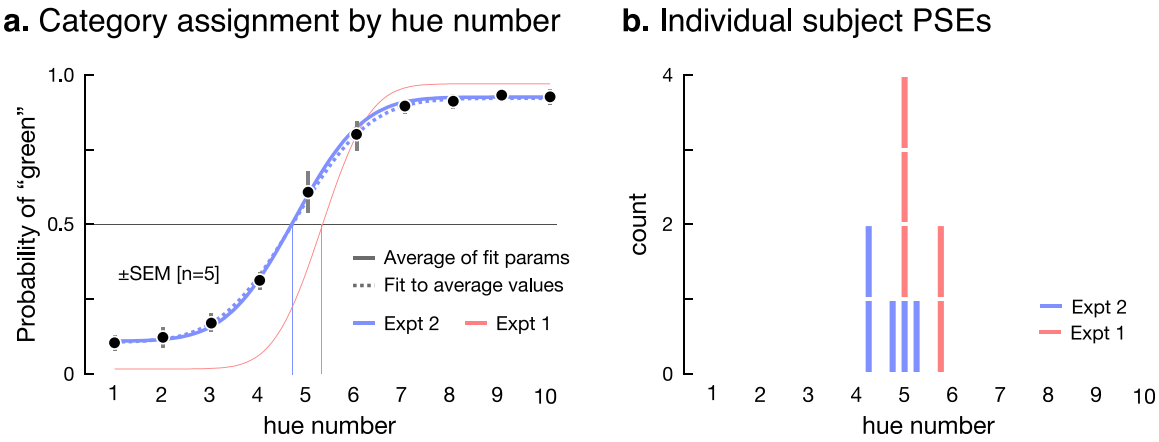

Figure S1. Psychophysical results in Experiment 2. (a) Blue–green categorization performance. Each point shows the average over subjects of the probability of a *green* response to each of our 10 colored checkerboards. The dashed blue line is the fit of a cumulative normal to these average data. The solid blue line is the cumulative normal obtained by averaging the mean and variance parameters of cumulative normal fit to each subject's data. The red line is the cumulative normal fit from Experiment 1, obtained from only the subset of five subjects who participated in Experiment 2. Error bars show ±1 *SEM*. (b) Histogram of PSEs obtained for individual subjects. Blue bars, Experiment 2; red bars, Experiment 1 for the same five subjects.

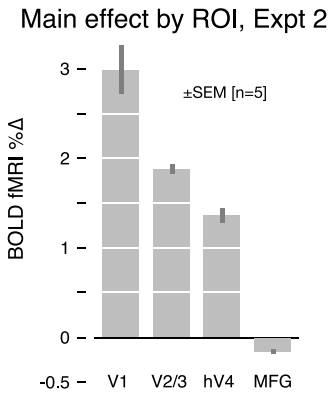

Figure S2. Main effect of color stimuli versus background on BOLD response in Experiment 2, shown by ROI.