Behavioral/Cognitive

Human V4 Activity Patterns Predict Behavioral Performance in Imagery of Object Color

[©]Michael M. Bannert^{1,2,3,4,5} and [©]Andreas Bartels^{1,2,3,4}

¹Vision and Cognition Laboratory, Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, 72076 Tübingen, Germany, ²Bernstein Center for Computational Neuroscience, 72076 Tübingen, Germany, ³Max Planck Institute for Biological Cybernetics, 72076 Tübingen, Germany, ⁴Department of Psychology, University of Tübingen, 72076 Tübingen, Germany, and ⁵International Max Planck Research School for Cognitive and Systems Neuroscience, 72076 Tübingen, Germany

Color is special among basic visual features in that it can form a defining part of objects that are engrained in our memory. Whereas most neuroimaging research on human color vision has focused on responses related to external stimulation, the present study investigated how sensory-driven color vision is linked to subjective color perception induced by object imagery. We recorded fMRI activity in male and female volunteers during viewing of abstract color stimuli that were red, green, or yellow in half of the runs. In the other half we asked them to produce mental images of colored, meaningful objects (such as tomato, grapes, banana) corresponding to the same three color categories. Although physically presented color could be decoded from all retinotopically mapped visual areas, only hV4 allowed predicting colors of imagined objects when classifiers were trained on responses to physical colors. Importantly, only neural signal in hV4 was predictive of behavioral performance in the color judgment task on a trial-by-trial basis. The commonality between neural representations of sensory-driven and imagined object color and the behavioral link to neural representations in hV4 identifies area hV4 as a perceptual hub linking externally triggered color vision with color in self-generated object imagery.

Key words: color vision; drift diffusion; fMRI; object imagery; pattern classification; reaction times

Significance Statement

Humans experience color not only when visually exploring the outside world, but also in the absence of visual input, for example when remembering, dreaming, and during imagery. It is not known where neural codes for sensory-driven and internally generated hue converge. In the current study we evoked matching subjective color percepts, one driven by physically presented color stimuli, the other by internally generated color imagery. This allowed us to identify area hV4 as the only site where neural codes of corresponding subjective color perception converged regardless of its origin. Color codes in hV4 also predicted behavioral performance in an imagery task, suggesting it forms a perceptual hub for color perception.

Introduction

Our ability to perceive the large variability in reflective properties of objects surrounding us is central to object recognition, scene segmentation and foraging (Mollon, 1989; Gegenfurtner and

Received Aug. 15, 2017; revised Feb. 5, 2018; accepted Feb. 8, 2018.

The authors declare no competing financial interests.

Correspondence should be addressed to either Dr. Michael M. Bannert or Andreas Bartels, Vision and Cognition Lab, Centre for Integrative Neuroscience, University of Tübingen, Otfried-Müller-Str. 25, 72076 Tübingen, Germany, E-mail: michael.bannert@tuebingen.mpg.de or andreas.bartels@tuebingen.mpg.de.

DOI:10.1523/JNEUROSCI.2307-17.2018

Copyright © 2018 the authors 0270-6474/18/383657-12\$15.00/0

Rieger, 2000; Dominy and Lucas, 2001). Not only can we perceive color effortlessly and automatically under most normal viewing conditions (Shevell and Kingdom, 2008; Brainard and Maloney, 2011; Foster, 2011), but also our memories, dreams, and thoughts can feature color just as naturally. Although much of visual cortex responds to chromatic input (Seidemann et al., 1999; Wandell et al., 1999; Liu and Wandell, 2005), the contributions of different neural representations to color perception is not clear (Conway, 2014; Zeki, 2015).

The brain shows some specialization for color as in the cytochrome oxidase-rich neural tissue in V1 (Livingstone and Hubel, 1984, 1988; Song et al., 2011) and V2 (Xiao et al., 2003; Sincich and Horton, 2005; Wang et al., 2007; Nasr et al., 2016), and V4's "glob cells" (Conway et al., 2007; Conway and Tsao, 2009; Tanigawa et al., 2010). Color processing extends to more anterior ventral regions (Brewer et al., 2005), including temporal cortex

Author contributions: M.M.B. and A.B. designed research; M.M.B. performed research; M.M.B. analyzed data; M.M.B. and A.B. wrote the paper.

This work was supported by the German Federal Ministry for Education and Research (BMBF) Grant number: FKZ 01GQ1002, the German Excellence Initiative of the German Research Foundation (DFG) Grant number: EXC307, the Max Planck Society, Germany, and the German Research Foundation (DFG): SFB 1233, Robust Vision: Inference Principles and Neural Mechanisms. TP 09.

(Lafer-Sousa and Conway, 2013; Lafer-Sousa et al., 2016), and also parietal cortex (Zeki and Stutters, 2013).

Several approaches aimed to isolate those aspects of sensorydriven neural activity that are related to perception. One involves identifying neural signals related to perceptual color constancy in varying physical illumination conditions. Corresponding illumination-induced shifts in neuronal tuning profiles have been found in V1 (MacEvoy and Paradiso, 2001; Wachtler et al., 2003) and in V4 (Zeki, 1980, 1983; Kusunoki et al., 2006), whereas doubleopponent tuning is thought to occur as early as V1 (Livingstone and Hubel, 1984; Johnson et al., 2001; Conway et al., 2002). In the human brain, engagement of color constancy was associated with increased activity in V1, V4, and in regions anterior to them (Bartels and Zeki, 2000; Barbur and Spang, 2008). For V1 as well as for V4 α , this activity has been shown to represent surface color regardless of illumination (Bannert and Bartels, 2017).

Another line of research related perceptual hue similarity to neural similarity measures. Although performing hue judgments involves V1 and V4 (Beauchamp et al., 1999), V4 activity reflected the similarity structure between hues and task-induced distortions of the perceptual hue plane (Brouwer and Heeger, 2009, 2013), even though selectivity for a range of perceptually relevant hues was found in all human visual areas from V1 to V4 (Goddard et al., 2010; Kuriki et al., 2015).

However, despite the wide use of decoding methods in color vision research (Brouwer and Heeger, 2009; Parkes et al., 2009; Seymour et al., 2009) and the interest in other forms of nonretinal color experience (Nunn et al., 2002; Hubbard et al., 2005; Gould van Praag et al., 2016), including imagery (Howard et al., 1998; Rich et al., 2006), no study has applied them to isolate neural activity that can be associated exclusively with perceptual content.

In this study we evoked converging subjective color percepts, one driven by physically presented color stimuli, the other by internally generated color imagery (Chang et al., 2013; Wantz et al., 2015). This allowed us to identify neural sites with converging color information that represent subjective color perception regardless of its origin. We implemented this approach by recording fMRI activity in two types of runs. In one run type, participants viewed abstract color stimuli that did not convey object information. In the other type, participants visually imagined objects, each of which was associated to a particular color. We then trained classifiers to discriminate between BOLD activity patterns elicited by veridical chromatic stimulation and tested them in predicting the color of the imagined objects. In addition, we related the trial-to-trial performance of classification to behavioral performance in a 1-back color judgment task. Our results identify color representations in hV4 as being shared between stimulus-driven and internally generated color vision and relate hV4 activity to behavioral performance in color imagery.

Materials and Methods

Participants

Nineteen volunteers participated in the experiment. They provided written informed consent before the first experimental session. One participant failed to complete the experiment. We thus analyzed data from the remaining volunteers (N = 18, 3 males, 15 females, age between 22–35 years, mean = 25.8 years). All participants had normal color vision as measured with Ishihara plates (Ishihara, 2011).

Stimuli

Objects of different colors and shapes as imagery templates

In the beginning of the experiment participants were familiarized with the images of the nine objects used in the imagery condition (Fig. 1). Each belonged to one of three color categories (red, green, and yellow). Each



Figure 1. Stimulus material. Stimuli used in the imagery task. Each object belonged to one of three color categories (yellow, red, green). To reduce confounds unrelated to object color, objects were approximately matched in shape (elongated round, pile-shaped) and semantic associations (all of them were fruits/vegetables). Before scanning and before each imagery fMRI run, participants practiced to remember the images of the nine natural objects (see Materials and Methods). During each imagery block, they imagined one of the objects, and mentally compared its color to that of the subsequent object once the word-cue appeared (Fig. 2).

object also belonged to one of three shape categories, such that each color could be presented in the form of three different shapes, or each shape could be presented in three different colors, leading to a 3×3 design. Shapes were elongated, round, or pile-shaped (Fig. 1). Furthermore all objects belonged to the same superordinate level category ("food") to minimize semantic confounds.

These objects were not shown during fMRI, where they were merely cued using words. The cue words in the imagery phase were presented in black letters in the center of a gray (154.1 cd/m^2) screen, at a text size of 0.77° of visual angle. Cue words were presented in German or English, depending on the participant's preference. All stimuli were presented with Psychtoolbox-3 (Kleiner et al., 2007).

Physical abstract color stimuli

In our "real-color runs" we presented abstract color stimuli of three different colors (red, green, and yellow) that would allow us to train classifiers on brain patterns evoked by different visually presented colors. The colors were presented in the form of concentric expanding rings that had mean chromaticities of x = 0.39, y = 0.35 in the red; x = 0.34, y = 0.41 in the green; and x = 0.41, y = 0.43 in the yellow conditions, respectively. The colored rings were shown at high and low psychophysically matched intensities (see below) on a gray background of medium luminance (154.1 cd/m²) to obtain luminance invariant color decoders.

Luminance values were determined for each color using the minimal flicker procedure (Kaiser, 1991) that required participants to adjust the luminance of a color stimulus presented against achromatic backgrounds at the high and low intensities (184.9 cd/m² and 151.3 cd/m², respectively) until the amount of perceived flicker was minimal. The color stimulus used for the minimal flicker method was a vertical rectangle presented in the display center and covering 3.28° vertically and 2.46° of visual angle horizontally. It was presented every second frame against the gray background while, in every other frame, the gray background was shown alone. The background intensity was either "high" or "low" (see above) to create luminance-matched color stimuli at both luminance levels. While lying in the scanner, participants adjusted the luminance of the color stimuli by button press in steps of 11.5 cd/m² and confirmed their adjustments by pressing another button. The six stimuli (3 hues, high- and low-intensity) were presented in random order. Mean luminance (SD in parentheses) for red, green, and yellow were 200.0 (9.1) cd/m², 189.6 (8.4) cd/m², and 179.0 (9.2) cd/m² for low-intensity stimuli and 244.2 (19.3) cd/m², 242.9 (11.3) cd/m², and 231.9 (11.1) cd/m² for high-intensity stimuli. In the main experiment, high and low



Figure 2. Experimental design. Trial sequence in real-color and imagery runs. In real-color runs (left) participants viewed concentric rings slowly drifting outward. The rings could be one of three colors: yellow, green, or red. Observers performed a detection task requiring them to press a button every time the luminance of the stimulus changed between high and low luminance. There could be 0, 1, or 2 target events per stimulus presentation (8.5 s, ITI = 1.5 s). In imagery runs (right), participants saw a word cue at the beginning of the trial for 1.5 s indicating which of the nine object images they had to imagine in the subsequent imagery phase (11.714 s). In addition to object imagery, participants had to perform a mental color comparison each time a new word cue appeared: they had to decide whether or not the color of the object they had to imagine in this trial matched the color of the object in the previous trial, and press one of two designated buttons accordingly (1-back same/different color judgment task).

luminance versions of the stimuli varied within real-color runs (see Paradigm and procedure) in a counterbalanced manner.

Note that our approach ensured that luminance across the three different colors were perceptually matched, and the use of two luminance levels for each color ensured that classifiers would be luminance invariant. Importantly though for our key finding, luminance was not relevant as this concerned cross-decoding between real color and imagery.

The concentric color rings were created by displaying a colored disc (radius: 8.61° of visual angle) that had its α (i.e., transparency) channel sinusoidally modulated as a function of eccentricity (2.16° visual angle cycle size). The rings drifted outward at a velocity of 2.47°/s.

Experimental design

Setup

Participants viewed the stimuli while lying supine in a scanner via a mirror fixed to the head coil. A projector (NEC PE401H) displayed the stimuli on a screen placed at the end of the scanner tube. The display was gamma-calibrated using a Photo Research PR-670 spectroradiometer (CalibrateMonSpd.m function from Psychtoolbox). Display size was 21.8° by 16.2° of visual angle at a resolution of 800 \times 600 pixels and a frame rate of 60 Hz.

Paradigm and procedure

Before the start of the scan, we familiarized the participants with the object images to ensure that they could correctly identify each object based on the word cue. They practiced the study task described below, which they performed later in the scanner before each imagery run.

The fMRI experiment consisted of six real-color and six imagery runs (Fig. 2) performed in alternation. Run type (real-color or imagery) of the first run was counterbalanced across participants. In real-color runs, participants viewed continuously expanding ring-shaped abstract color stimuli in blocks of 8.5 s and responded with a button press whenever the color changed luminance for a brief period of time (0.3 s). Participants received feedback about their task performance at the end of each run for motivation before the next run started. A total of 216 color blocks (36 per run) entered the analysis. Blocks of color/luminance combinations (3 colors \times 2 luminances = 6 conditions) were presented in a pseudorandomized sequence, which ensured that each color/luminance pair was preceded by every pair an equal number of times (Brooks, 2012). Each real-color run started with a block that was identical to the last one of the

preceding real-color run to keep the back-matched sequence intact. This initial block was not included in the analysis.

Before the start of each imagery run, participants performed a study task to memorize the objects they had to visualize during the measurement. This study task required participants to distinguish the correct object from among three distractor objects of the same basic level category. The participants had to perform this practice task until they successfully completed one trial, in which they correctly identified each of the nine objects, before the fMRI experiment was resumed. After initial training outside the scanner, this study criterion was usually reached after one or two trials.

For the imagery task, participants were instructed to imagine the objects as if they were seeing them on the screen. During imagery runs, participants fixated on a small circle in the middle of the screen. A cue word appeared at the beginning of each imagery block for 1.5 s indicating which of the nine objects to mentally visualize in the subsequent imagery block. The imagery block lasted 11.714 s after which a new cue word of the subsequent trial appeared. Upon the appearance of a new cue word, participants performed a 1-back color judgment task, which required them to indicate if the object referred to by the new cue and the object they had been imaging previously were the same color. We instructed our participants to reply as quickly and accurately as possible. They had to make a decision while the cue word was still on the screen. At the end of each run they received feedback about reaction times and errors for motivation. As in real-color runs, each of the nine object conditions was preceded equally often by every other condition in a first-order historymatched pseudorandomized sequence (Brooks, 2012). The first blocks in every run were used only to keep the presentation sequence intact across runs. In total 162 imagery blocks (27 per run) entered the analysis.

Retinotopic mapping and ROI definition

Each volunteer participated in a separate retinotopic mapping session to identify their visual areas V1, V2, V3, hV4, VO1, LO1, and LO2. We chose these areas because they had previously been shown to be involved in the processing of shape and color (Brewer et al., 2005; Larsson and Heeger, 2006; Brouwer and Heeger, 2009; Seymour et al., 2010). We used standard retinotopic mapping procedures to identify reversals in the angle map of the visual field representations that delineate the boundaries between these areas on the cortical surface (Sereno et al., 1994; Wandell and Winawer, 2011). Participants viewed a contrast-reversing checker-

board through a wedge-shaped aperture. Because the visual field representations are compressed at large eccentricities on the cortical surface (cortical magnification), the check sizes increased logarithmically with eccentricity. The aperture subtended the entire screen within a 90° angle at all eccentricities from the fixation dot. The wedge rotated at a period of 55.64 s for a total of 10 cycles per run. The rotation direction alternated between the four mapping runs.

fMRI scan details

We measured BOLD activity with a 64-channel head coil at 3T magnetic field strength (Siemens Prisma) using 56 slices oriented axially but slightly tilted in parallel with the AC-PC line. The sampling volume covered almost the whole brain with a slice thickness of 2 mm and no gap between slices. In plane resolution was 96×96 , yielding an isotropic voxel size of 2 mm. We used a fourfold GRAPPA accelerated parallel imaging sequence (GRAPPA factor 2) to measure T2*-weighted functional images. Repetition time (TR) and echo time (TE) were 0.87 s and 30 ms, respectively, with a flip angle of 57°. Anatomical images with an isotropic voxel size of 1 mm were measured using a T1-weighted MP-RAGE ADNI sequence and magnetic field inhomogeneities were measured with a gradient echo field map.

fMRI data preprocessing

The first 11 functional images recorded per run were discarded to allow the MRI signal to reach equilibrium. Functional data were realigned to correct for head motion and unwarped using the estimated field map, slice time corrected, and coregistered to the anatomical image. Finally the data were normalized to MNI space using a segmentation-based normalization of the anatomical image. No smoothing was applied to the images from the main experiment. We used SPM8 (http://www.fil.ion.ucl.ac.uk/ spm) for preprocessing. The data from the retinotopic mapping session underwent the same preprocessing up to coregistration in SPM8. The resulting images were then further preprocessed with FreeSurfer (http:// surfer.nmr.mgh.harvard.edu/), which involved smoothing them with a 4 mm Gaussian kernel. Individual cortical surfaces for all participants were obtained using FreeSurfer's recon-all pipeline.

Statistical analysis

Aims and strategy

The aim of this study was to examine how color of imagined objects is represented in relation to sensory-driven color. We estimated vectors of fMRI responses using a standard GLM approach and then performed pattern classification analyses on these data.

Our analysis strategy followed two main steps: in a first step we trained classifiers to distinguish the three colors an observer was seeing using only data from real-color runs (real-color-to-real-color classification). To verify this training procedure, we used cross-validation leaving out data from a different run in every iteration to obtain an unbiased accuracy estimate for the classifier. In a second step, we trained classifiers on responses from all real-color runs but this time tested them on responses from the imagery runs (real-color-to-imagined color classification). Crucially, this analysis tested for commonalities between the representation of color shown as abstract color rings and the color of imagined objects.

Furthermore, we wanted to know whether participants actually performed the object imagery task instead of merely keeping the object color in mind. We therefore performed another classification analysis to decode the shape of the imagined object (imagined-to-imagined shape classification). Because the three shapes were roughly matched across color categories, we could thus train classifiers to predict the shape of the object ("elongated", "round", or "pile-shaped") that they imagined on a given imagery trial while balancing the number of colors across shape categories. If the shape of the imagined object could be decoded, this would indicate that participants indeed activated a neural representation of shape even though the judgment task did not require participants to mentally represent this stimulus dimension.

Control analysis for shape decoding: comparing shape coding against word coding with representational similarity analysis This control analysis served to shed further light on our shape decoding analysis, and is not directly relevant to our main analysis of crossdecoding between real and imagined color. Shape decoding could in principle be mediated by neural signals related to the visual word cue, because before each imagery phase a unique word appeared on screen indicating which object to imagine. Although we were careful to model the BOLD response in the imagery phase with regressors placed at the offset of the cue presentation (see Pattern estimation), we additionally conducted the present control analysis using representational similarity analysis (RSA; Kriegeskorte et al., 2008; Kriegeskorte, 2011; Kriegeskorte and Kievit, 2013; Nili et al., 2014) to directly compare how well the data reflect either shape information or word cue properties. We computed the representational dissimilarities between the responses evoked by the nine imagery objects [this time without recursive feature elimination (RFE) so as not to bias the analysis toward shape encoding] as the Mahalanobis distances between pairs of activity patterns cross-validated across the six imagery runs (Walther et al., 2016).

The representational dissimilarity matrix (RDM) of pairwise distances obtained in this way were rank-correlated (Spearman's ρ) with two model RDMs: the shape RDM contained zeros for pairs belonging to the same shape category and ones otherwise; the control RDM contained the difference between word pairs in terms of the number of character insertions, deletions, and substitutions required to change one word into the other ("Levenshtein" model). The Levenshtein distance (also called "edit distance") is a common measure of string similarity (Duda et al., 2001, p. 418) that simultaneously takes into account similarity with respect to important properties like, e.g., word length or letter sequences shared between words. The Levenshtein RDM was rescaled so that its entries ranged from 0 to 1. English or German word RDMs were used for individual datasets depending on the language in which the cues were presented to the participant. Levenshtein RDMs, however, were strongly rank-correlated between languages (0.7528).

Pattern estimation

We modeled the unsmoothed voxel time series with one boxcar regressor per trial using SPM8. In the real-color runs, the onset times of each color block served as regressor onsets. To avoid contamination with visual processing related to the displayed cue, we chose the cue offset times as regressor onsets in imagery runs. All regressors were shifted 5 s forward in time to account for the hemodynamic lag. Realignment parameters regressed out any linear dependence between head motion and voxel time series. We estimated one β parameter in every voxel for each of the 216 real-color and 162 imagery trials across all runs. Estimates from different voxels were combined to form vectors of brain responses. In every voxel, the time series of β estimates was quadratically detrended by removing, in every run, the fit of a second order polynomial from the original data to filter out low-frequency noise. Each residual time series was z-scored for each run separately. To make our analysis more robust against outliers, we set all values with a difference of >2 SD from the mean to -2 and 2, respectively.

Classification details

We used linear discriminant analysis (LDA) classifiers for pattern classification using the Princeton MVPA Toolbox (https://github.com/ PrincetonUniversity/princeton-mvpa-toolbox) and inhouse MATLAB code. Due to the low number of samples and high dimensionality of the dataset we used a shrinkage estimator for the covariance matrix to ensure that it remained nonsingular (Ledoit and Wolf, 2004). Additionally, we applied RFE (De Martino et al., 2008) on training data only to select the set of voxels that optimally distinguished between the categories to be classified. The optimal set of voxels was then used to fit the classifier to the entire training set and to validate it on the test set, which was not part of the voxel selection procedure. RFE determined the optimal voxel set by repeatedly training LDA classifiers on part of the training set (i.e., leaving out 1 run each time) and testing it on the remaining part of the training set (i.e., the withheld run) to obtain an accuracy score. This procedure was repeated 15 times while each time dropping those 15% of the voxels from the classification whose coefficients varied the least across discriminant functions and hence were least discriminative of the category to be predicted.

Statistical inference

We used permutation tests to evaluate the statistical significance of our classifications results. For every time we trained a classifier to discriminate between fMRI patterns, we refit new classifiers after randomly permuting the labels in the training set 10^3 times. The reasoning for this is that under the null hypothesis of no association between fMRI patterns and category membership (e.g., color category) the labels can be randomly reassigned to fMRI vectors without changing the expected classification accuracies. We used the 10^3 classification accuracies from each participant to obtain a null distribution of mean accuracies at the group level expected under the null hypothesis (including the accuracy that was actually observed using the unpermuted dataset). From these null distributions, *p* values for a one-tailed test can be calculated as the number of values in the distribution that exceeded the observed accuracies divided by the number of permutations.

Because we examined classification accuracies from several ROIs, we needed to correct for multiple comparisons. We controlled the familywise error (FWE) by constructing a common null distribution for all ROIs by taking the maximum value across ROI group means in each permutation step while making sure that the same label permutations were used in every ROI (Nichols and Holmes, 2002). The resulting null distribution was then used to calculate FWE-corrected *p* values.

In our RSA we tested whether correlations between model RDMs and brain-derived RDMs were significantly larger than 0 by bootstrapping the error distribution around the observed correlation (10^4 iterations). For one-tailed testing, the probability of observing a deviation from 0 at least as large as the observed one was divided by two. We controlled for multiple comparisons by applying a false discovery rate threshold of q = 0.05.

Linking behavior to neural coding using drift diffusion models

Drift diffusion model. To enhance the mental representation of color during object imagery, we had instructed our participants to indicate whether the color of the object imagined in the current trial was identical with that of the subsequent trial. The task was performed as soon as the word cue for the next object appeared. To quantify behavioral task performance using a unified model taking all behavioral measures into account, we fitted reaction times (RTs) and errors using hierarchical drift diffusion models (HDDM; Wiecki et al., 2013). Drift diffusion models provide a principled way to integrate RTs and errors in a single value called a drift rate. Drift rate quantifies how quickly evidence is accumulated in a decision-making process and thus indicates how easily a task is performed. Higher values mean shorter RTs and fewer errors. An important advantage of HDDM is that it estimates model parameters by directly modeling the dependencies between model parameters across subjects in a hierarchical Bayesian framework for improved sensitivity. For model details and parameter priors see Wiecki et al. (2013, p. 3).

We used HDDM to probe whether neural information as decoded in the real-color-to-imagined fMRI classification related to the behavioral performance in the imagery task. The hypothesis was that when participants were well engaged in the imagery task, this would improve the behavioral performance as well as the signal of the neural representation of the imagined object color. It follows that behavioral performance is better when color classifiers made correct compared with incorrect predictions for imagery trials. To test this, we fit a single HDDM to behavioral data that assumed separate drift rates for correctly and incorrectly classified trials. According to the above hypothesis, drift rates for correctly classified trials were expected to be higher than on incorrectly classified trials.

Bayesian inference. We then used Bayesian parameter estimation to calculate the posterior probability of the drift rate being indeed higher for correct than incorrect trials. The relationship between brain signal and behavior was computed separately for the fMRI signal preceding the behavioral judgment or for the signal following behavioral judgment (see Fig. 5A). In one model different drift rates were assumed depending on whether the imagery pattern immediately following the color judgment was classified correctly ("post-judgment model") whereas the second model assumed different drift rates for correct versus incorrect classifications of the pattern that preceded the color judgment ("pre-judgment

model"). These two patterns were thought to be relevant since the task required a comparison about the color of the previous with the following objects. An additional model that checked dependence on classification two trials before the color judgment (data not shown) was also included for comparison.

We performed Markov chain Monte Carlo (MCMC) sampling to approximate the posterior distribution over model parameters given the data using 10 chains. Each of them drew 10⁴ samples plus additional 10³ for burn-in. Using MCMC sampling, we obtained a posterior probability distribution for the difference between drift rates for correctly and incorrectly decoded trials. This distribution represents how likely each value of this difference is, given the behavioral data and the classifier predictions. In Bayesian parameter estimation, the posterior probability that the difference is larger than 0 is then given by the probability mass occupying the interval above 0.

Statistical inference was inherently corrected for multiple comparisons by false discovery rate: when conducting statistical inference on the basis of posterior probabilities in a family of tests, a cutoff of 0.95 automatically enforces a corresponding upper bound of 5% on the false discovery rate (Friston and Penny, 2003). The reason is that, with each statistical decision, there is a 5% probability that the null hypothesis has been falsely rejected, leading to an expected fraction of false-positives of 5% for the whole test family. Note that only trials with key presses made within the 1.5 s response window were included in this analysis, i.e., while the cue was still on the screen.

Results

Real-color decoding

First we validated our methods by testing whether classifiers could predict which color a person was seeing from the multivariate pattern of fMRI responses when participants viewed the colored ring stimuli. Figure 3A illustrates that real color could be decoded from fMRI activity in all ROIs we studied (p < 0.001, FWE corrected), replicating previous findings (Brouwer and Heeger, 2009; Seymour et al., 2009). With a chance level of 33%, classification accuracies averaged across participants were as follows: 51.4% in V1 [SD = 5.3%, Cohen's *d* = 3.38, 95% CI (48.9, 53.9)], 51.4% [SD = 6.6%, Cohen's d = 2.72, 95% CI (48.3, 54.5)] in V2, 49.3% [SD = 7.6%, Cohen's d = 2.1,95% CI (45.8,52.8)] in V3,49% [SD = 7.1%, Cohen's *d* = 2.2, 95% CI (45.7, 52.3)] in hV4, 42.8% [SD = 5.5%, Cohen's d = 1.74,95% CI (40.3,45.4)] in VO1, and 43.6% [SD = 5.8%, Cohen's d = 1.76, 95% CI (40.9, 46.3)] in LO1, 41.4% in LO2 [SD = 6.7%, Cohen's *d* = 1.2, 95% CI (38.3, 44.5)], all *p* = 0.001, FWE corrected. This shows that it is possible to construct a classifier to reliably decode sensory-driven color.

Predicting imagined object color

Our main hypothesis was that the neural representations of imagined object color overlap with those of veridical color perception. The crucial test for this was to train classifiers to distinguish between sensory-driven colors and to test them on the imagery trials. If sensory-driven and imagined object color representations overlap in a given ROI, one would expect classification accuracies above chance level. As can be seen from Figure 3*B*, the color of imagined objects could indeed be decoded successfully from area hV4 [M = 36.2%, SD = 2.6%, p = 0.005, FWE corrected, Cohen's d = 1.08, 95% CI (35.0, 37.4)], but in no other brain region. This means that the color-specific patterns of fMRI activity elicited by object imagery resembled those measured during sensory-driven color vision in hV4 only.

To test for potential confounding behavioral effects on decoding accuracies, we compared RTs and error rates in the 1-back same/different task between color conditions but did not find any differences for RTs ($F_{(2,34)} = 0.005$, p = 0.9947) or errors ($F_{(2,34)} =$ 0.5, p = 0.6106, one-way repeated-measures ANOVA, respectively). Our analysis thus did not provide evidence for cross-decoding being driven by behavioral differences between color conditions.

Mean decoding accuracies in the remaining areas were 32.5% [SD = 3.6%, p = 1, Cohen's d = -0.22, 95% CI (30.9, 34.2)] in V1, 34.2% [SD = 4.2%, p = 0.667, Cohen's d = 0.21, 95% CI (32.3, 36.2)] in V2, 33.5% [SD = 3.0%, p = 0.971, Cohen's d = 0.06, 95% CI (32.1, 34.9)] in V3, 33.2% [SD = 4.2%, p = 0.993, Cohen's d = -0.04, 95% CI (31.2, 35.1)] in VO1, 32.1% [SD = 3.6%, p = 1, Cohen's d = -0.35, 95% CI (30.4, 33.7)] in LO1, and 34.1% [SD = 3.2%, p = 0.757, Cohen's d = 0.25, 95% CI (32.7, 35.6)] in LO2 (all FWE corrected).

Because all visual areas, including V1 and V2, allowed for decoding of physically presented color, the failure to cross-decode imagined color cannot be explained by poor signal quality relative to hV4. As shown in the previous section, decoding accuracies for sensory-driven color tended to be slightly larger in V1 and V2 than in hV4. Post hoc *t* tests revealed that this difference, however, marginally failed to reach significance [V1: $t_{(17)} = 1.4189$, p = 0.087, 95% CI (-1.18, 6.01); V2: $t_{(17)} = 1.619$, p = 0.0619, 95% CI (-0.73, 5.51), each one-tailed and uncorrected].

Decoding the shape of imagined objects

Next we examined whether our participants did as instructed imagine the objects as a whole, i.e., including object shape, even though this was irrelevant to the color judgment task. The objects in this experiment had been chosen such that there were three identical types of shape in every color category: elongated, round, and pile-shaped. If observers imagined the entire objects instead of just retaining a (possibly non-pictorial) representation of that object's color in mind, we would expect shape information also to be represented in the neural signal evoked during

object imagery. We tested this by training classifiers on imagery responses to discriminate between the three shape categories (regardless of color) and testing them on fMRI responses from a run that was not part of the training set (leave-one-run-out cross-validation). As shown in Figure 4, the imagined object shape could be successfully decoded from areas V2 [M = 38.1%, SD = 4.8%, p = 0.001, FWE corrected, Cohen's d = 1.01, 95% CI (35.9, 40.3)], V3 [M = 37.4%, SD = 5.9%, p = 0.001, FWE corrected, Cohen's d = 0.68, 95% CI (34.6, 40.1)], and LO2 [M = 35.7%, SD = 5.4%, p = 0.01, FWE corrected, Cohen's d = 0.44, 95% CI (33.2, 38.2)]. This is consistent with the interpretation that participants' object imagery on average encompassed also the representation of object shape although it was irrelevant to the color judgment task. This result once again demonstrates that sufficient



Figure 3. Color decoding results. Real-color-to-real-color and real-color-to-imagined color decoding. *A*, Classifiers were trained to distinguish between the three color categories based on the responses in each ROI to the color stimuli. Classifier performance was cross-validated leaving out one of the six runs for testing on each iteration to obtain an average accuracy. In all ROIs classifiers could predict the color of the stimulus that participants were viewing significantly above chance. *B*, Color classifiers were trained on the whole set of fMRI responses to color stimuli to predict which color observers were seeing. The learned classifiers were then used to predict on a trial-by-trial basis the color of the objects that participants were imagining in the imagery runs of the experiment. Permutation tests showed that the color of the imagined objects could be decoded significantly above chance only from activity patterns in area hV4. *A*, *B*, Horizontal and vertical bars represent group means and two-tailed 95% confidence intervals, respectively. Chance level was 1/3. **p < 0.01, FWE corrected, one-tailed.

signal was present also in imagery trials in early as well as high-level regions to decode imagery content.

Because the imagery phase for each object was preceded by a unique cue word, we tested for possible confounding effects on shape decoding. Specifically, we used RSA (Nili et al., 2014) to directly probe how well the similarity structure between fMRI responses to the nine imagined objects was predicted by model RDMs that either captured shape information or cue word information (Levenshtein distance of word pairs, see Materials and Methods, Experimental design).

One-tailed bootstrap tests showed that the group average rank correlations between the shape model and brain-derived RDMs were significantly positive in areas V3 [M = 0.1039, SD = 0.1518, p = 0.0023, FDR adjusted, 95% CI (0.03, 0.17)] and LO2 [M =



Figure 4. Shape decoding results. Classifiers were trained to distinguish between the three shapes (elongated round, pile-shaped) and tested on data from one of the six imagery runs that was excluded from the training procedure in a sixfold leave-one-run-out cross-validation scheme. The shape property of objects was orthogonal to the color feature dimension (Fig. 1). Mean decoding accuracies were significantly above chance in areas V2, V3, and L02 according to permutation tests. Horizontal and vertical bars represent group means and two-tailed 95% confidence intervals, respectively. Chance level was 1/3. *p < 0.05, **p < 0.01, FWE corrected, one-tailed.

0.0803, SD = 0.1618, p = 0.027, FDR adjusted, 95% CI (0.01, 0.15)] but not in V2 [M = 0.0531, SD = 0.1474, p = 0.0612, uncorrected, 95% CI (-0.02, 0.12)]. The Levenshtein model conversely was significantly correlated with brain-derived RDMs in V2 [M = 0.1546, SD = 0.3154, p = 0.0336, FDR adjusted, 95% CI (0.0, 0.3)] and V3 [M = 0.1301, SD = 0.308, p = 0.0321, FDR adjusted, 95% CI (-0.02, 0.27)]. Crucially and in contrast to the shape model, however, there was no significant correlation in LO2 [M = 0.0712, SD = 0.2302, p = 0.092, uncorrected, 95% CI (-0.04, 0.18)].

Our RSA approach thus shows that, although similarity structures of BOLD responses could to some degree indeed be explained by word similarity (V2) or both shape and word similarity (V3), the similarity structure of LO2 activity correlated with shape information exclusively. This confirms our conclusion that participants represented shape information (i.e., imagined objects), which was irrelevant to the behavioral color judgment task.

Brain-based drift diffusion modeling

We identified hV4 as a visual area where the color of imagined objects could be decoded using color classifiers trained on sensorydriven color responses. It is unclear, however, whether the neuronal signal underlying this observation played a behaviorally relevant role for imagery of object color. We therefore sought to investigate the relationship between the predictions of the brain signal classifier and our participants' behavior in the color judgment task, i.e., errors and RTs.

Mean RT across the group was 0.853 s (SD = 0.075 s) and mean error rate was 5.3% (SD = 4.5%). Mean RT in correct and in error trials were 0.853 s (SD = 0.075 s) and 0.866 s (SD = 0.134s), respectively. Trials without a button press during the 1.5 sresponse window were excluded from the analysis.

We fitted HDDMs to our participants' behavioral data (errors and RTs) and tested whether the drift rate in these models differed for imagery trials in which neural signal led to correct or incorrect decoding of object color. Several convergence diagnostics including visual inspection of trace plots, autocorrelation between samples at different lags, and Gelman-Rubin statistic calculations (for all parameters, R < 1.02) showed that the chains had converged to their stationary distributions. The distributions obtained through MCMC sampling approximate the posterior probability density function only when convergence is reached. Here, the chains could hence be combined and used for Bayesian inference.

Bayesian inference showed that drift rates were higher when color decoders correctly predicted the color of an imagined object in the trial that immediately preceded the color judgment. This finding held true only for area hV4. The posterior probability of a higher drift rate on correct trials given the data **D** was $p(v_c > v_i | \mathbf{D}) =$ 0.9802 (Fig. 5B, pre-judgment model). This means that the drift rate on correctly predicted trials was larger than on incorrectly predicted trials with probability 0.9802 (inherently FDR-corrected; see Materials and Methods), given the behavioral data and classifier outputs. It demonstrates that observers performed the color judgment more easily when the ob-

ject color could be correctly predicted from hV4 activity measured in the imagery blocks immediately preceding the color judgment task. Accordingly, RTs were shorter following an imagery block that was correctly classified (M = 0.844 s, SD = 0.077) than following incorrectly classified ones [M = 0.86 s, SD = 0.077, $t_{(17)} = 2.4711$, p = 0.0243, Cohen's d = 0.5825, 95% CI (0.002, 0.029), one-tailed paired t test]. Error rates did not differ significantly between correct (M = 4.7%, SD = 4.6%) and incorrect classifications [M = 5.6%, SD = 4.8%, $t_{17} = 1.0132$, p = 0.3252, Cohen's d = 0.2388, 95% CI (-1, 3), one-tailed paired t test]. The posterior probabilities for the models fit to data in the other ROIs did not reach the 0.95 threshold [V1: $p(v_c > v_i | D) = 0.2714$; V2: $p(v_c > v_i | D) = 0.219$; V3: $p(v_c > v_i | D) = 0.7298$; VO1: $p(v_c > v_i | D) = 0.8296$; LO1: $p(v_c > v_i | D) = 0.2277$; LO2: $p(v_c > v_i | D) = 0.4801$].

When considering the classifier prediction of the imagery trial immediately following the color judgment, the posterior probability for hV4 did not exceed the 0.95 threshold and dropped to $p(v_c > v_i | \mathbf{D}) = 0.9275$ (Fig. 5*B*, post-judgment model). Likewise, drift rates were not larger on correctly than incorrectly decoded trials in any of the other ROIs [V1: $p(v_c > v_i | \mathbf{D}) = 0.6$; V2: $p(v_c > v_i | \mathbf{D}) = 0.0663$; V3: $p(v_c > v_i | \mathbf{D}) = 0.7447$; VO1: $p(v_c > v_i | \mathbf{D}) = 0.4414$; LO1: $p(v_c > v_i | \mathbf{D}) = 0.3165$; LO2: $p(v_c > v_i | \mathbf{D}) = 0.2921$]. Also when assuming different drift rates depending on classifier accuracies two trials before the color judgment, the posterior probabilities in none of the ROIs exceeded the threshold of 0.95 [V1: $p(v_c > v_i | \mathbf{D}) = 0.5894$; V2: $p(v_c > v_i | \mathbf{D}) = 0.2907$; V3: $p(v_c > v_i | \mathbf{D}) = 0.2214$; hV4: $p(v_c > v_i | \mathbf{D}) = 0.8709$; VO1: $p(v_c > v_i | \mathbf{D}) = 0.3923$; LO1: $p(v_c > v_i | \mathbf{D}) = 0.1151$; LO2: $p(v_c > v_i | \mathbf{D}) = 0.2416$].

The correlation between behavioral performance with the correctness of the classifier predictions in hV4 suggests that the neural information represented in hV4 on the color of the imagined object was behaviorally relevant for correct and rapid task execution in the following trial.



Figure 5. Brain-based drift diffusion modeling. HDDMs were used to study the relationship between the information content in ROIs and performance in the color judgment task. *A*, Behavioral data (RTs and errors) on trial*j* were modeled with different drift rates depending on whether the classifier predicted the correct or the incorrect label either for the response pattern y_{j-1} on the previous trial*j*. (pre-judgment model), which was recorded before the behavioral judgment, or the response pattern y_j on the same trial *j*, which was recorded after the behavioral judgment (post-judgment model). As depicted on the right, higher drift rates mean faster response times and fewer errors. *B*, Posterior probability distribution over the difference v_{diff} between drift rates v_c on correctly classified and v_i on incorrectly classified trials for the two models in area hV4. The posterior probability of v_c being larger than v_i was 98.02% in the pre-judgment model, i.e., when different drift rates were assumed depending on the classification of the trial preceding the behavioral judgment. However, when instead classifier correctness for the imagery block following the behavioral judgment was taken into account (post-judgment model), the posterior probability dropped to 92.75%. See main text for values of all ROIs.

Discussion

We examined neural representations of color during sensory stimulation and during imagery of object color. This allowed us to identify neural sites representing subjective color experience regardless of its origin. Classifiers trained to distinguish between sensory-driven colors could predict the color of the object our participant was imagining based on the activity only in hV4. Importantly, the quality of the neural code only in this area predicted behavioral performance in the color imagery task. The results suggest that hV4 forms a perceptual hub for color perception.

Common neural representations for color imagery and sensory-driven color vision

The similarity of activity patterns in object imagery and sensorydriven color vision is consistent with its role in color perception (Lueck et al., 1989; Bartels and Zeki, 2000; Bouvier and Engel, 2006; Brouwer and Heeger, 2009, 2013; Conway and Tsao, 2009). The perceptual relevance of color-selective activity in hV4 is corroborated by the psychophysiological link we discovered using brain-based behavioral modeling. Observers performed better in their task when object color could successfully be decoded from hV4 activity immediately before the behavioral decision. This provides an important additional clue to the perceptual nature of this neural signal because perception evolved to guide behavior (Friston, 2010; Purves et al., 2011; Hoffman et al., 2015).

To the extent that imagery and working memory rely on the same mechanisms (Keogh and Pearson, 2011; Albers et al., 2013), our results are thus consistent with theories ascribing a central role to perceptual representations in visual short-term memory and imagery (Finke, 1980; Pasternak and Greenlee, 2005; D'Esposito and Postle, 2015). Memory for color has been shown to modulate neuronal excitability and to elicit sustained firing of V4 neurons (Ferrera et al., 1994; Motter, 1994). Interestingly, simultaneous recordings from V4 and prefrontal cortex showed phase-locking of local field potentials (LFPs) to predict behavioral accuracy in working memory for colored objects (Liebe et al., 2012). Because LFP correlates well with BOLD activity (Logothetis, 2008), this may suggest a neural mechanism underlying the relationship between hV4 activity and task performance observed in our study.

The present findings hence extend early studies reporting increased fMRI responses to sensory-driven and imagined color in V4 (Rich et al., 2006) or, instead, anterior to it (Howard et al., 1998) in important ways: we show that activity in hV4 was selective for the precise content of internally generated color experience, thereby signaling the color of the imagined object. This distinguishes it from color representations in other visual areas like V1, which we could only show to represent sensory-driven color vision. Such dissociations may reflect that color is a nonunitary attribute linked to multiple perceptual primitives (Mausfeld, 2003) and is hence represented at multiple processing stages.

An alternative view could entail a component of categorical coding of color in hV4 with tighter chromatic color codes in earlier regions. If this were the case, small deviations of imagined color from the trained real color stimuli would degrade cross-classification in early but not late stages of color representation. More research is hence needed to illuminate the relationship between color perception, hue, and category encoding. A starting point is given by prior studies that found frontal lobe regions to be involved in color category encoding (Bird et al., 2014), or V4 activity patterns to predict behavior in color categorization (Brouwer and Heeger, 2013).

Coding for color and spatial features in V4

Given that V4 is involved in the processing of object properties such as shape (Kobatake and Tanaka, 1994; Pasupathy and Connor, 2002; Dumoulin and Hess, 2007; Bushnell and Pasupathy, 2012) and texture (Kohler et al., 2016), it is plausible that there is an overlap in the coding of color in perception and object imagery. This is because the internal generation of object percepts requires a unified representation of several different object-related features (shape, texture, etc.), which may engage particularly those neural representations affording a suitable degree of feature binding with color (Seymour et al., 2009, 2010, 2016).

Explicit and implicit processes in color vision

Our findings are potentially important in illuminating a difference in neural mechanisms between processes related to active imagery as opposed to those related to completion, association or error-correction during perception. V1 and higher areas may play different roles when viewed from a process-based perspective that distinguishes between top-down signals that can be characterized as either explicit and voluntary or as implicit and involuntary (Albright, 2012; Pearson and Westbrook, 2015).

In a previous study we had presented participants grayscale photographs of objects that are typically associated to a given object-specific color, such as bananas, strawberries, etc. Participants performed a motion discrimination task while viewing grayscale object images. Classifiers trained on sensory-driven color responses predicted the unperceived memory color of the grayscale objects in V1 but not in extrastriate regions (Bannert and Bartels, 2013). The contrast to the current findings suggests that active imagery and automatic color-association related to objects recruit fundamentally distinct mechanisms: conscious and attentive object imagery engages color encoding in hV4, whereas automatic (and likely unconscious) association of memory colors to objects engages V1. This account is consistent with research suggesting that color-selectivity of BOLD signals in early areas seem to reflect implicit top-down influences on color processing (Amano et al., 2016), whereas in higher areas this tends to apply rather for explicit effects (Brouwer and Heeger, 2013; Vandenbroucke et al., 2016).

Color imagery and visual attention

From a process-based viewpoint, our findings therefore have to be discussed especially with respect to visual attention as an explicit form of top-down influence. There are two reasons for this: first, attention plays a central role in the binding of object features (Treisman, 1988; Humphreys, 2016) and, second, because attention is thought to rest on similar cognitive top-down mechanisms as working memory (Kastner and Ungerleider, 2000; Gazzaley and Nobre, 2012; D'Esposito and Postle, 2015). Both properties make it hence likely for object imagery to be accomplished by attending to internal object feature representations. Modulation of V4 activity by attention to color is well established by electrophysiology (Moran and Desimone, 1985; McAdams and Maunsell, 2000) and fMRI (Bartels and Zeki, 2000; Saenz et al., 2002; Brouwer and Heeger, 2013). Feature-based attention can flexibly change the spatial tuning of V4 neurons along stimulus dimensions depending on task demands in visual search (David et al., 2008). It is plausible that such task-dependent changes may be expressed as changes in presynaptic integration processes, which can be detected with fMRI and may depend on (top-down) input from other brain regions (Liebe et al., 2011).

The role of early visual areas

The fact that color classification did not generalize from sensorydriven color to imagined object color in V1 or V2 does not imply that such an effect could not be obtained with more sensitive methods or even that those areas do not partake in imagery of object color. One imaging study conducted at 7T field strength showed that imagined pieces of art could be identified from V1 and V2 activity (Naselaris et al., 2015), but it did not identify the unique contributions of different visual features (and none of them was color). We do not believe, however, that poor sensitivity accounted for the absence of color imagery signals in early visual cortex in the present study because physically presented color could be predicted at least as accurately from fMRI patterns in V1 or V2 as in hV4. We interpret the difference that we observed between early visual areas and hV4 therefore as reflecting the increase in sensitivity to top-down processing like attention for higher visual areas, which may give rise to a perception/imagery gradient in the visual cortex (Lee et al., 2012).

It has been argued that the involvement of V1 in imagery is task-dependent such that imagery tasks requiring more detailed information (Pearson et al., 2015, p. 596), may activate more low-level visual features as well. The fact that we did not find effects in V1 suggests that imagery of object-associated color relies primarily on extrastriate coding, possibly fostered by differences in object-color associations in categorical terms rather than fine detail. Accordingly, a decoding study involving a task requiring working memory and discrimination of only very small changes in color saturation identified V1 as hue-encoding site (Serences et al., 2009). In contrast, the present study decoded categorical hue differences between red, green, and yellow and involved imagery of highly distinct objects rather than working memory of fine detail. Note though that direct comparisons between working memory and imagery are not unconditionally permissible because participants may have pursued different cognitive strategies to meet task demands (Keogh and Pearson, 2011).

Conclusion

Our experiment directly related sensory driven color representations in the brain to those generated internally through imagery. The results show that in the entire visual cortex, those representations converge in hV4. We hence identified hV4 as a neural site bridging the domains of sensory-driven and imagined object color. The behavioral relevance of activity in this area highlights its role in generating color percepts, be they externally triggered or a product of our minds.

References

- Albers AM, Kok P, Toni I, Dijkerman HC, de Lange FP (2013) Shared representations for working memory and mental imagery in early visual cortex. Curr Biol 23:1427–1431. CrossRef Medline
- Albright TD (2012) On the perception of probable things: neural substrates of associative memory, imagery, and perception. Neuron 74:227–245. CrossRef Medline
- Amano K, Shibata K, Kawato M, Sasaki Y, Watanabe T (2016) Learning to associate orientation with color in early visual areas by associative decoded fMRI neurofeedback. Curr Biol 26:1861–1866. CrossRef Medline
- Bannert MM, Bartels A (2013) Decoding the yellow of a gray banana. Curr Biol 23:2268–2272. CrossRef Medline
- Bannert MM, Bartels A (2017) Invariance of surface color representations across illuminant changes in the human cortex. Neuroimage 158:356– 370. CrossRef Medline
- Barbur JL, Spang K (2008) Colour constancy and conscious perception of changes of illuminant. Neuropsychologia 46:853–863. CrossRef Medline
- Bartels A, Zeki S (2000) The architecture of the colour centre in the human visual brain: new results and a review. Eur J Neurosci 12:172–193. CrossRef Medline
- Beauchamp MS, Haxby JV, Jennings JE, DeYoe EA (1999) An fMRI version of the Farnsworth–Munsell 100-hue test reveals multiple color-selective areas in human ventral occipitotemporal cortex. Cereb Cortex 9:257–263. CrossRef Medline
- Bird CM, Berens SC, Horner AJ, Franklin A (2014) Categorical encoding of color in the brain. Proc Natl Acad Sci U S A 111:4590–4595. CrossRef Medline
- Bouvier SE, Engel SA (2006) Behavioral deficits and cortical damage loci in cerebral achromatopsia. Cereb Cortex 16:183–191. CrossRef Medline
- Brainard DH, Maloney LT (2011) Surface color perception and equivalent illumination models. J Vis 11(5):1 1–18. CrossRef Medline
- Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat Neurosci 8:1102– 1109. CrossRef Medline
- Brooks JL (2012) Counterbalancing for serial order carryover effects in experimental condition orders. Psychol Methods 17:600–614. CrossRef Medline
- Brouwer GJ, Heeger DJ (2009) Decoding and reconstructing color from responses in human visual cortex. J Neurosci 29:13992–14003. CrossRef Medline
- Brouwer GJ, Heeger DJ (2013) Categorical clustering of the neural representation of color. J Neurosci 33:15454–15465. CrossRef Medline
- Bushnell BN, Pasupathy A (2012) Shape encoding consistency across colors in primate V4. J Neurophysiol 108:1299–1308. CrossRef Medline

- Chang S, Lewis DE, Pearson J (2013) The functional effects of color perception and color imagery. J Vis 13(10):4 1–10. CrossRef Medline
- Conway BR (2014) Color signals through dorsal and ventral visual pathways. Vis Neurosci 31:197–209. CrossRef Medline
- Conway BR, Tsao DY (2009) Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. Proc Natl Acad Sci U S A 106:18034–18039. CrossRef Medline
- Conway BR, Hubel DH, Livingstone MS (2002) Color contrast in macaque V1. Cereb Cortex 12:915–925. CrossRef Medline
- Conway BR, Moeller S, Tsao DY (2007) Specialized color modules in macaque extrastriate cortex. Neuron 56:560–573. CrossRef Medline
- David SV, Hayden BY, Mazer JA, Gallant JL (2008) Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. Neuron 59:509–521. CrossRef Medline
- De Martino F, Valente G, Staeren N, Ashburner J, Goebel R, Formisano E (2008) Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. Neuroimage 43:44–58. CrossRef Medline
- D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. Annu Rev Psychol 66:115–142. CrossRef Medline
- Dominy NJ, Lucas PW (2001) Ecological importance of trichromatic vision to primates. Nature 410:363–366. CrossRef Medline
- Duda RO, Hart PE, Stork DG (2001) Pattern classification, Ed 2. New York: Wiley.
- Dumoulin SO, Hess RF (2007) Cortical specialization for concentric shape processing. Vision Res 47:1608–1613. CrossRef Medline
- Ferrera VP, Rudolph KK, Maunsell JH (1994) Responses of neurons in the parietal and temporal visual pathways during a motion task. J Neurosci 14:6171–6186. Medline
- Finke RA (1980) Levels of equivalence in imagery and perception. Psychol Rev 87:113–132. CrossRef Medline
- Foster DH (2011) Color constancy. Vision Res 51:674–700. CrossRef Medline
- Friston K (2010) The free-energy principle: a unified brain theory? Nat Rev Neurosci 11:127–138. CrossRef Medline
- Friston KJ, Penny W (2003) Posterior probability maps and SPMs. Neuroimage 19:1240–1249. CrossRef Medline
- Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. Trends Cogn Sci 16:129–135. CrossRef Medline
- Gegenfurtner KR, Rieger J (2000) Sensory and cognitive contributions of color to the recognition of natural scenes. Curr Biol 10:805–808. CrossRef Medline
- Goddard E, Mannion DJ, McDonald JS, Solomon SG, Clifford CW (2010) Combination of subcortical color channels in human visual cortex. J Vis 10(5):25 1–17. CrossRef Medline
- Gould van Praag CD, Garfinkel S, Ward J, Bor D, Seth AK (2016) Automaticity and localisation of concurrents predicts colour area activity in grapheme-colour synaesthesia. Neuropsychologia 88:5–14. CrossRef Medline
- Hoffman DD, Singh M, Prakash C (2015) The interface theory of perception. Psychon Bull Rev 22:1480–1506. CrossRef Medline
- Howard RJ, Ffytche DH, Barnes J, McKeefry D, Ha Y, Woodruff PW, Bullmore ET, Simmons A, Williams SC, David AS, Brammer M (1998) The functional anatomy of imagining and perceiving colour. Neuroreport 9:1019–1023. CrossRef Medline
- Hubbard EM, Arman AC, Ramachandran VS, Boynton GM (2005) Individual differences among grapheme-color synesthetes: brain-behavior correlations. Neuron 45:975–985. CrossRef Medline
- Humphreys GW (2016) Feature confirmation in object perception: feature integration theory 26 years on from the Treisman Bartlett lecture. Q J Exp Psychol (Hove) 69:1910–1940. CrossRef Medline
- Ishihara S (2011) Ishihara's tests for colour deficiency, 38 plates. Tokyo: Kanehara Trading.
- Johnson EN, Hawken MJ, Shapley R (2001) The spatial transformation of color in the primary visual cortex of the macaque monkey. Nat Neurosci 4:409–416. CrossRef Medline
- Kaiser PK (1991) Flicker as a function of wavelength and heterochromatic flicker photometry. In: Limits of vision (Kulikowski JJ, Walsh V, Murray IJ, eds), pp 171–190. Basingstoke, UK: MacMillan.

Bannert and Bartels • Neural Color Representations in Visual Imagery

- Kastner S, Ungerleider LG (2000) Mechanisms of visual attention in the human cortex. Annu Rev Neurosci 23:315–341. CrossRef Medline
- Keogh R, Pearson J (2011) Mental imagery and visual working memory. PLoS One 6:e29221. CrossRef Medline
- Kleiner M, Brainard DH, Pelli DG (2007) What's new in Psychtoolbox-3? [Abstract] Perception 36 (ECVP Abstract Supplement). CrossRef
- Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J Neurophysiol 71:856–867. CrossRef Medline
- Kohler PJ, Clarke A, Yakovleva A, Liu Y, Norcia AM (2016) Representation of maximally regular textures in human visual cortex. J Neurosci 36:714– 729. CrossRef Medline
- Kriegeskorte N (2011) Pattern-information analysis: from stimulus decoding to computational-model testing. Neuroimage 56:411–421. CrossRef Medline
- Kriegeskorte N, Kievit RA (2013) Representational geometry: integrating cognition, computation, and the brain. Trends Cogn Sci 17:401–412. CrossRef Medline
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60:1126–1141. CrossRef Medline
- Kuriki I, Sun P, Ueno K, Tanaka K, Cheng K (2015) Hue selectivity in human visual cortex revealed by functional magnetic resonance imaging. Cereb Cortex 25:4869–4884. CrossRef Medline
- Kusunoki M, Moutoussis K, Zeki S (2006) Effect of background colors on the tuning of color-selective cells in monkey area V4. J Neurophysiol 95:3047–3059. CrossRef Medline
- Lafer-Sousa R, Conway BR (2013) Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. Nat Neurosci 16: 1870–1878. CrossRef Medline
- Lafer-Sousa R, Conway BR, Kanwisher NG (2016) Color-biased regions of the ventral visual pathway lie between face- and place-selective regions in humans, as in macaques. J Neurosci 36:1682–1697. CrossRef Medline
- Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. J Neurosci 26:13128–13142. CrossRef Medline
- Ledoit O, Wolf M (2004) Honey, I shrunk the sample covariance matrix. J Portf Manag 30:110–119. CrossRef
- Lee SH, Kravitz DJ, Baker CI (2012) Disentangling visual imagery and perception of real-world objects. Neuroimage 59:4064–4073. CrossRef Medline
- Liebe S, Logothetis NK, Rainer G (2011) Dissociable effects of natural image structure and color on LFP and spiking activity in the lateral prefrontal cortex and extrastriate visual area V4. J Neurosci 31:10215–10227. CrossRef Medline
- Liebe S, Hoerzer GM, Logothetis NK, Rainer G (2012) Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. Nat Neurosci 15:456–462. CrossRef Medline
- Liu J, Wandell BA (2005) Specializations for chromatic and temporal signals in human visual cortex. J Neurosci 25:3459–3468. CrossRef Medline
- Livingstone MS, Hubel DH (1984) Anatomy and physiology of a color system in the primate visual cortex. J Neurosci 4:309–356. Medline
- Livingstone M, Hubel D (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 240:740–749. CrossRef Medline
- Logothetis NK (2008) What we can do and what we cannot do with fMRI. Nature 453:869–878. CrossRef Medline
- Lueck CJ, Zeki S, Friston KJ, Deiber MP, Cope P, Cunningham VJ, Lammertsma AA, Kennard C, Frackowiak RS (1989) The colour centre in the cerebral cortex of man. Nature 340:386–389. CrossRef Medline
- MacEvoy SP, Paradiso MA (2001) Lightness constancy in primary visual cortex. Proc Natl Acad Sci U S A 98:8827–8831. CrossRef Medline
- Mausfeld R (2003) Colour as part of the format of different perceptual primitives: the dual coding of colour. In: Colour Perception: mind and the physical world (Mausfeld R, Heyer D, eds), pp 381–430. Oxford, UK: Oxford UP.
- McAdams CJ, Maunsell JH (2000) Attention to both space and feature modulates neuronal responses in macaque area V4. J Neurophysiol 83: 1751–1755. CrossRef Medline
- Mollon JD (1989) "Tho' she kneel'd in that place where they grew": the uses and origins of primate colour vision. J Exp Biol 146:21–38. Medline

- Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784. CrossRef Medline
- Motter BC (1994) Neural correlates of attentive selection for color or luminance in extrastriate area V4. J Neurosci 14:2178–2189. Medline
- Naselaris T, Olman CA, Stansbury DE, Ugurbil K, Gallant JL (2015) A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. Neuroimage 105:215–228. CrossRef Medline
- Nasr S, Polimeni JR, Tootell RB (2016) Interdigitated color- and disparityselective columns within human visual cortical areas V2 and V3. J Neurosci 36:1841–1857. CrossRef Medline
- Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 15:1–25. CrossRef Medline
- Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N (2014) A toolbox for representational similarity analysis. PLoS Comput Biol 10:e1003553. CrossRef Medline
- Nunn JA, Gregory LJ, Brammer M, Williams SC, Parslow DM, Morgan MJ, Morris RG, Bullmore ET, Baron-Cohen S, Gray JA (2002) Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. Nat Neurosci 5:371–375. CrossRef Medline
- Parkes LM, Marsman JB, Oxley DC, Goulermas JY, Wuerger SM (2009) Multivoxel fMRI analysis of color tuning in human primary visual cortex. J Vis 9(1):1 1–13. CrossRef Medline
- Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. Nat Rev Neurosci 6:97–107. CrossRef Medline
- Pasupathy A, Connor CE (2002) Population coding of shape in area V4. Nat Neurosci 5:1332–1338. CrossRef Medline
- Pearson J, Westbrook F (2015) Phantom perception: voluntary and involuntary nonretinal vision. Trends Cogn Sci 19:278-284. CrossRef Medline
- Pearson J, Naselaris T, Holmes EA, Kosslyn SM (2015) Mental imagery: functional mechanisms and clinical applications. Trends Cogn Sci 19: 590–602. CrossRef Medline
- Purves D, Wojtach WT, Lotto RB (2011) Understanding vision in wholly empirical terms. Proc Natl Acad Sci U S A 108:15588–15595. CrossRef Medline
- Rich AN, Williams MA, Puce A, Syngeniotis A, Howard MA, McGlone F, Mattingley JB (2006) Neural correlates of imagined and synaesthetic colours. Neuropsychologia 44:2918–2925. CrossRef Medline
- Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human visual cortex. Nat Neurosci 5:631–632. CrossRef Medline
- Seidemann E, Poirson AB, Wandell BA, Newsome WT (1999) Color signals in area MT of the macaque monkey. Neuron 24:911–917. CrossRef Medline
- Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. Psychol Sci 20:207–214. CrossRef Medline
- Sereno MI, McDonald CT, Allman JM (1994) Analysis of retinotopic maps in extrastriate cortex. Cereb Cortex 4:601–620. CrossRef Medline
- Seymour K, Clifford CW, Logothetis NK, Bartels A (2009) The coding of color, motion, and their conjunction in the human visual cortex. Curr Biol 19:177–183. CrossRef Medline
- Seymour K, Clifford CW, Logothetis NK, Bartels A (2010) Coding and binding of color and form in visual cortex. Cereb Cortex 20:1946–1954. CrossRef Medline
- Seymour KJ, Williams MA, Rich AN (2016) The representation of color across the human visual cortex: distinguishing chromatic signals contributing to object form versus surface color. Cereb Cortex 26:1997–2005. CrossRef Medline
- Shevell SK, Kingdom FA (2008) Color in complex scenes. Annu Rev Psychol 59:143–166. CrossRef Medline
- Sincich LC, Horton JC (2005) The circuitry of V1 and V2: integration of color, form, and motion. Annu Rev Neurosci 28:303–326. CrossRef Medline
- Song JH, Rowland J, McPeek RM, Wade AR (2011) Attentional modulation of fMRI responses in human V1 is consistent with distinct spatial maps for chromatically defined orientation and contrast. J Neurosci 31:12900– 12905. CrossRef Medline
- Tanigawa H, Lu HD, Roe AW (2010) Functional organization for color and orientation in macaque V4. Nat Neurosci 13:1542–1548. CrossRef Medline
- Treisman A (1988) Features and objects: the fourteenth Bartlett memorial lecture. Q J Exp Psychol A 40:201–237. CrossRef Medline
- Vandenbroucke ARE, Fahrenfort JJ, Meuwese JDI, Scholte HS, Lamme VAF (2016) Prior knowledge about objects determines neural color represen-

tation in human visual cortex. Cereb Cortex 26:1401–1408. CrossRef Medline

- Wachtler T, Sejnowski TJ, Albright TD (2003) Representation of color stimuli in awake macaque primary visual cortex. Neuron 37:681–691. CrossRef Medline
- Walther A, Nili H, Ejaz N, Alink A, Kriegeskorte N, Diedrichsen J (2016) Reliability of dissimilarity measures for multi-voxel pattern analysis. Neuroimage 137:188–200. CrossRef Medline
- Wandell BA, Winawer J (2011) Imaging retinotopic maps in the human brain. Vision Res 51:718–737. CrossRef Medline
- Wandell BA, Poirson AB, Newsome WT, Baseler HA, Boynton GM, Huk A, Gandhi S, Sharpe LT (1999) Color signals in human motion-selective cortex. Neuron 24:901–909. CrossRef Medline
- Wang Y, Xiao Y, Felleman DJ (2007) V2 thin stripes contain spatially organized representations of achromatic luminance change. Cereb Cortex 17:116–129. CrossRef Medline

Wantz AL, Borst G, Mast FW, Lobmaier JS (2015) Colors in mind: a novel

paradigm to investigate pure color imagery. J Exp Psychol Learn Mem Cogn 41:1152–1161. CrossRef Medline

- Wiecki TV, Sofer I, Frank MJ (2013) HDDM: hierarchical bayesian estimation of the drift-diffusion model in python. Front Neuroinform 7:14. CrossRef Medline
- Xiao Y, Wang Y, Felleman DJ (2003) A spatially organized representation of colour in macaque cortical area V2. Nature 421:535–539. CrossRef Medline
- Zeki S (1980) The representation of colours in the cerebral cortex. Nature 284:412–418. CrossRef Medline
- Zeki S (1983) Colour coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelengths and colours. Neuroscience 9:741–765. CrossRef
- Zeki S (2015) Area V5: a microcosm of the visual brain. Front Integr Neurosci 9:21. CrossRef Medline
- Zeki S, Stutters J (2013) Functional specialization and generalization for grouping of stimuli based on colour and motion. Neuroimage 73:156–166. CrossRef Medline