

## Prior Knowledge about Objects Determines Neural Color Representation in Human Visual Cortex

A. R. E. Vandenbroucke<sup>1,2</sup>, J. J. Fahrenfort<sup>3</sup>, J. D. I. Meuwese<sup>1,2</sup>, H. S. Scholte<sup>1,2</sup> and V. A. F. Lamme<sup>1,2</sup>

<sup>1</sup>Cognitive Neuroscience Group, Department of Psychology, University of Amsterdam, 1018 XA Amsterdam, the Netherlands,

<sup>2</sup>Cognitive Science Center Amsterdam, University of Amsterdam, 1018 WS Amsterdam, the Netherlands and <sup>3</sup>Department of Cognitive Psychology, Vrije Universiteit, 1081 BT Amsterdam, the Netherlands

Address correspondence to Annelinde R. E. Vandenbroucke, Department of Psychology, University of Amsterdam, Weesperplein 4, 1018 XA Amsterdam, the Netherlands. Email: vandenbroucke.work@gmail.com

**To create subjective experience, our brain must translate physical stimulus input by incorporating prior knowledge and expectations. For example, we perceive color and not wavelength information, and this in part depends on our past experience with colored objects (Hansen et al. 2006; Mitterer and de Ruiter 2008). Here, we investigated the influence of object knowledge on the neural substrates underlying subjective color vision. In a functional magnetic resonance imaging experiment, human subjects viewed a color that lay midway between red and green (ambiguous with respect to its distance from red and green) presented on either typical red (e.g., tomato), typical green (e.g., clover), or semantically meaningless (nonsense) objects. Using decoding techniques, we could predict whether subjects viewed the ambiguous color on typical red or typical green objects based on the neural response of veridical red and green. This shift of neural response for the ambiguous color did not occur for nonsense objects. The modulation of neural responses was observed in visual areas (V3, V4, VO1, lateral occipital complex) involved in color and object processing, as well as frontal areas. This demonstrates that object memory influences wavelength information relatively early in the human visual system to produce subjective color vision.**

**Keywords:** color, fMRI, MVPA, object knowledge, subjective perception

### Introduction

As the brain processes incoming information, visual representations become detached from the low-level properties of stimulus input: The visual world is instantly interpreted to match our beliefs and expectations. Object knowledge, for example, influences the colors we perceive (Hansen et al. 2006; Mitterer and de Ruiter 2008; Witzel et al. 2011): We expect bananas to be yellow and carrots to be orange.

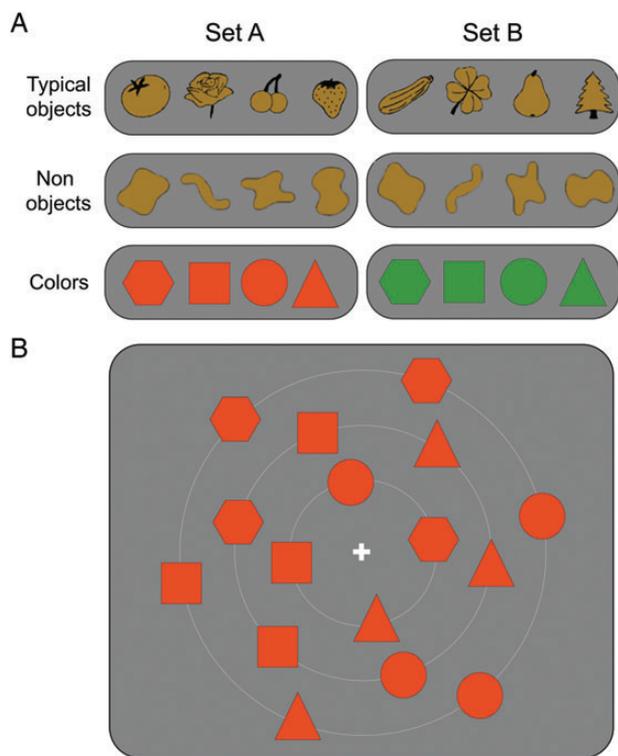
Behavioral studies have shown that when subjects have to indicate whether they perceive a color that lies exactly between orange and yellow as either orange or yellow, they more often categorize the ambiguous color as yellow when presented on typical yellow objects (e.g., a banana), and as orange when presented on typical orange objects (e.g., a carrot; Mitterer and de Ruiter 2008). Moreover, when participants have to adjust the color of a fruit picture such that it appears achromatic, they tend to overcompensate in the direction of the opponent hue (Hansen et al. 2006; Witzel et al. 2011). When a similar color is presented on a scrambled pattern, participants do not overcompensate. This suggests that color categorization is not only influenced by semantic knowledge, but that visual perception itself is adjusted according to object-color associations.

Which neural substrates underlie the incorporation of prior knowledge about objects to create subjective color perception, however, remains unknown. Previous research has shown that

striate and extrastriate areas V1 and V2 are color selective; however, cells in these areas are primarily color-opponent and luminance-dependent (Zeki 1983; Brouwer and Heeger 2009; Shapley and Hawken 2011). Area V4 and visual areas anterior to V4 (VO1), on the other hand, have been shown to be involved in color constancy (Zeki and Marini 1998; Heywood and Kenridge 2003; but for critical reviews see Gegenfurtner and Kiper 2003; Shapley and Hawken 2011) and are suggested to respond according to perceptual color space rather than to low-level color properties (Brouwer and Heeger 2009). Area V3 is functionally grouped with area V4, as opposed to with area V1 and V2. Possibly, mid and higher level areas beyond V4 serve to combine color perception with memory for objects, thereby influencing neural responses to color in lower level areas and creating subjective color experience (Shapley and Hawken 2011).

In this study, we investigated which neural substrates underlie the effect that object knowledge has on our subjective color experience. Specifically, we examined whether early visual areas (V1, V2, V3, V4) merely represent bottom-up color attributes or whether they are influenced by prior knowledge. Using functional magnetic resonance imaging (fMRI), we determined whether the neural representation of a color that lies midway between red and green (ambiguous with respect to its distance from red and green) can be shifted toward red when presented on typical red objects (typical-red: tomato, strawberry, rose, cherry) and toward green when presented on typical green objects (typical-green: pine tree, clover, pear, zucchini) (Fig. 1A, top). Subjects were presented with blocks of ambiguously colored typical-red and blocks of ambiguously colored typical-green objects. As a control, subjects viewed 2 sets of semantically meaningless (nonsense set A and nonsense set B; Fig. 1A, middle) objects that were filled in with the same ambiguous color. For each condition, the representation of the ambiguous color was compared with veridical red and veridical green: runs containing blocks of red and blocks of green geometrical shapes (Fig. 1A, bottom) were presented after the ambiguously colored object runs. Thus, although the same ambiguous color was presented in 4 different object conditions, a shift toward red or green was only expected in the typical-red and typical-green object conditions.

Because mean activity change per condition (as used in classic univariate analyses) can be insensitive to differences between the processing of colors (Brouwer and Heeger 2009; Parkes et al. 2009; Seymour et al. 2009), we used multivoxel pattern analysis (MVPA) to characterize the neural responses underlying red, green, and the ambiguous color. We applied a support vector machine (SVM) algorithm in 4 functionally



**Figure 1.** Stimuli and task design. (A) Stimuli used for the typical object runs, nonsense object (nonobjects) runs and color runs. Note that colors will appear differently on different screens and on printouts. (B) Example of the positioning of the 16 objects around the fixation cross. The same rules applied to the typical and nonsense objects. Objects rotated around the fixation cross on the path of 3 circles (white lines, not visible in experiments: radius of 2.9°, 4.7°, and 6.9°, containing 4, 6, and 6 objects). The positioning of the objects was determined randomly at the start of each block, with the constraint that each object was placed once in the inner circle. Note that relative sizes are adjusted for display purposes.

defined visual regions of interest (ROIs; V1–V4) to determine whether we could predict which typical object set subjects were viewing based on the multivoxel pattern for veridical red and green. We hypothesized that if early visual areas represent physical wavelength information and are not influenced by object knowledge, the activity patterns underlying the ambiguous color should not be classifiable as either red or green (neither when presented on the typical objects nor on the ambiguously colored nonsense objects). On the other hand, if color representations in early visual areas are influenced by object knowledge, activity should be affected by object-color associations, resulting in a shift in the representation of the ambiguous color toward red or green for the typical objects, but not for the nonsense objects. In addition to investigating these functionally defined ROIs, we employed the same decoding technique in a whole-brain searchlight analysis (Kriegeskorte et al. 2006) to investigate whether areas other than early visual cortex were engaged in a shift in color representation.

## Materials and Methods

### fMRI Acquisition

Ten subjects (1 male, mean age = 23.5, SD = 4.5) participated in this experiment voluntarily or for monetary reward. All subjects had normal or corrected-to-normal vision and were tested on color vision using the Ishihara color blindness test. The study was approved by the local

Ethics Committee of the University of Amsterdam and subjects were screened on risk factors precluding participation from MRI experiments.

Scanning was performed on a 3T Philips TX Achieva MRI scanner at the Spinoza Center in Amsterdam. A high-resolution  $T_1$ -weighted anatomical image (TR 8.17 ms; TE 3.74 ms; FOV 240 × 220 × 188) was recorded for each subject. Functional MRI was recorded using a sagittally oriented gradient-echo, echo-planar pulse sequence (TR 2000 ms; TE 27.63 ms; FA 76°; 37 slices with interleaved acquisition; voxel size 2.5 × 2.5 × 3 mm; 80 × 80 matrix; FOV 200 × 200 × 122). Stimuli were back-projected on a 61 × 36 cm LCD screen using Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA) and viewed through a mirror attached to the head coil.

### Stimuli

To determine the red and green neural responses for each subject, 16 geometrical shapes (4 squares, 4 triangles, 4 circles, 4 hexagons; mean object size = 1.3° × 1.3°, SD = 0.1° × 0.1°; Fig. 1A, bottom) were simultaneously presented surrounding a white fixation cross (as in Fig. 1B). To optimize the level of activity in lower visual areas, the objects rotated around the white fixation cross during a 12-s presentation (Seymour et al. 2009). The objects were placed on 3 circles (radius: 2.9°, 4.7°, and 6.9°; movement speed: 1.5°/s, -2.5°/s, 3.6°/s) containing 4, 6, and 6 objects. At the start of each trial, the position of each object was randomly determined, with the constraint that one of each of the 4 different objects was placed on the inner circle (see Fig. 1). Red and green were adjusted to be isoluminant (red: CIE  $L^*$  = 55,  $a^*$  = 60,  $b^*$  = 54; green: CIE  $L^*$  = 55,  $a^*$  = -40,  $b^*$  = 38; both 65 cd/m<sup>2</sup>) with respect to the gray background (CIE  $L^*$  = 55,  $a^*$  = 0,  $b^*$  = 0; 65 cd/m<sup>2</sup>) based on their CIE  $L^*a^*b^*$  color values. The ambiguous color was chosen to lie exactly in between red and green (CIE  $L^*$  = 55,  $a^*$  = 10,  $b^*$  = 46; 65 cd/m<sup>2</sup>). Although some might label this color yellow, orange, or brown, we refer to it as ambiguous, since it is ambiguous with respect to the labels red and green. To map the responses for the ambiguous color in a specific object-color association, 16 typically green objects (typical-green: 4 pine trees, 4 clovers, 4 pears, 4 zucchini; M obj. size = 1.3° × 1.3°, SD = 0.2° × 0.1°) and 16 typically red objects (typical-red: 4 tomatoes, 4 strawberries, 4 roses, 4 cherries; M obj. size = 1.2° × 1.4°, SD = 0.2° × 0.2°) were filled in with the ambiguous color (Fig. 1A, top). These objects were line drawings taken from various sources on the internet, which we modified such that the typical-green and typical-red object sets had the same amount of colored and black pixels. As a control, 16 nonsense objects (4 different figures computed in Matlab using sine/cosine transforms; M obj. size = 1.4° × 1.4°, SD = 0.2° × 0.2°) were filled in with the ambiguous color and presented in 2 conditions, with the figures in the one condition (set A) rotated 90° compared with the other condition (set B) (Fig. 1A, middle).

We specifically chose to use line drawings instead of grayscale photographs. Although the memory color effect might be larger with grayscale photographs (Witzel et al. 2011), we wanted to minimize differences in local contrast or other image statistics between objects. Although the overall luminance can be adjusted to be equal in grayscale photographs, the differences in edges and local luminance cannot be controlled (think for example of a tree with many edges because of its leaves and a tomato that has only outer edges). This might especially influence the areas that are involved in processing local contrasts such as V1. To minimize these differences between stimuli, we used line drawings containing only ambiguously colored and black pixels.

### fMRI Procedure and Task

Subjects performed 12–16 runs depending on whether they performed one or two scanning sessions. Four subjects underwent 1 scanning session in which 4 typical object runs, 4 nonsense object runs and 4 color runs were recorded. Two subjects started with the typical object runs and 2 subjects started with the nonsense object runs. All subjects ended with the 4 color runs to make sure that there was no effect of seeing veridical red or veridical green on the perception of the ambiguous color. For example, subjects might (un)consciously introduce a bias when mapping the ambiguous color to either red or green, if they had already seen veridical red and green prior to the ambiguous color. Such an association might have overtaken any real object-color

associations, thus potentially creating a type II error. To prevent this, we showed the color runs at the end of the experiment. Six subjects underwent 2 scanning sessions of 8 runs. In the first session, 4 typical object runs and 4 color runs were recorded and in the second session, 4 nonsense object runs and 4 color runs were recorded. In this session, the color runs were recorded again in the second session to reduce variance that could be caused by imperfect registration or different blood oxygen level-dependent (BOLD) dynamics between sessions.

In each fMRI run, 2 conditions (color run: red and green; or typical run: typical-red and typical-green; or nonsense run: nonsense objects set A and set B) were presented for 10 blocks each (resulting in 20 blocks per run). One block lasted 16 s containing a 4-s rest period. Each run started with a 16-s rest period and ended with a 20-s rest period (total run time 356 s, 178 TR). Blocks were presented pseudo-randomly per run, and block presentation was counterbalanced between runs. During the whole run, subjects had to fixate on the fixation cross and press a button when the cross turned into a circle (3 times per 12-s presentation period, timed randomly). We used eye-tracking (Eyelink-1000, SR Research) to make sure that subjects properly fixated throughout each run.

### Behavioral Task

On a different day after the scanning session(s), subjects performed a behavioral task in which the subjective perception of the ambiguously colored stimuli was investigated. Subjects rated 7 colors on a red-to-green continuum as either red or green (see Fig. 4A, *x*-axis). The 7 colors (ranging from CIE  $L^* = 55$ ,  $a^* = 25$ ,  $b^* = 52$  to CIE  $L^* = 55$ ,  $a^* = -10$ ,  $b^* = 40$  in steps of  $a^* = 5$  and  $b^* = 2$ ) were presented on the objects used in the fMRI runs (Fig. 1A). Each stimulus was presented once a block and there were 5 blocks (4 objects  $\times$  4 conditions  $\times$  7 colors  $\times$  5 blocks = 560 trials total). This gave a total of 20 observations per color per condition (4 objects  $\times$  5 repetitions). A colored noise mask consisting of pixels that were all randomly colored was presented in between each presentation to prevent spillover effects of the previous trial.

### Region-of-Interest Localization

We functionally defined 5 visual regions of interest (ROIs: V1, V2, V3, V4, as well as all regions combined). Each subject completed 1–4 polar mappings and 1–2 eccentricity mappings. We used standard techniques to draw out early visual areas on the basis of these mappings (see Supplementary Fig. 1 and Wandell and Winawer 2011). For polar angle mapping, a checkerboard (red–green, flickering at 8 Hz) wedge rotated around fixation (clockwise or counterclockwise; complete revolution in 32 s; 8 repetitions) and, for eccentricity mapping, a checkerboard ring (red–green, flickering at 8 Hz) expanded from center to periphery (or vice versa; complete revolution in 32 s; 8 repetitions). During runs, subjects fixated at the center while detecting blue squares presented in the red–green checkerboard stimuli to keep their attention with the stimuli and maximize the visual response.

### Multivoxel Pattern Analyses

Data were analyzed using Brainvoyager QX 2.2 (Brain Innovation, Maastricht, the Netherlands, Goebel et al. 2006) and Matlab 2010 (MathWorks, Inc., Natick, MA, USA). Functional scans were slice-time corrected, motion corrected, spatially smoothed with a Gaussian of 2 mm FWHM and high-pass filtered at 0.01 Hz. All functional scans were first aligned to the functional scan that was recorded closest in time to the  $T_1$ -weighted anatomical image, and co-registered to the anatomical image that was transformed to Talairach space using an ACPC transform (Talairach and Tournoux 1988). The transformations that were necessary to co-register the EPI sequences to the subject-specific anatomical image and subsequently to the normalized anatomical image were concatenated, and were therefore handled as one single transformation. Moreover, the transformation from subject space to Talairach space is a piecewise linear transformation, and does not introduce nonlinearity (Poldrack et al. 2011). We normalized all subject-specific anatomical images to Talairach space to be able to compare our individual ROI analyses with the group-level searchlight analysis. Importantly, because the anatomical and functional data of each subject underwent the same subject-specific transformation into Talairach

space, the relative location of visual areas was maintained. All visual ROIs were functionally defined using subject-specific retinotopy and eccentricity mappings.

Data in the functional color and object runs were *z*-transformed. For each stimulus block, the 4 volumes (8 s) that corresponded to the peak of the BOLD-response were averaged [the peak of the BOLD-response was calculated for each subject separately by cross-correlating the *z*-transformed data of the training runs for the combined ROI (com) with a Gaussian ( $\alpha$  2.5)]. This created a specific voxel pattern for each subject and each stimulus presentation, separately for each ROI.

We used the Princeton MVPA toolbox (<http://code.google.com/p/princeton-mvpa-toolbox>) in combination with a SVM from the Bioinformatics toolbox to train the classifiers. To validate the MVPA technique on our dataset, we first performed a within-category classification in which we attempted to predict which color was presented (red or green), which typical object set was presented (typical-red or typical-green), and which nonsense object set was presented (set A vs. set B) using a leave-1-out procedure. In these analyses, 3 runs were used for training and the fourth run was used to test whether the presentations in this run could be correctly classified. All combinations of runs were once used for training, using each run once for testing. Data for these 4 iterations were averaged, yielding a classification score (proportion correct) for each subject and each ROI.

We used a permutation test to establish whether the 3 classifications (color, typical objects, and nonsense objects) significantly deviated from chance. The permutation test consisted of iterating the following 2 steps: 1) For each subject, the training labels were randomly permuted, such that the relationship between the labels and the imaging data were lost. This created a training set that relied on the same subject-specific data, which were used to create a random classifier. The same permutation was used for all subjects on any given iteration (i.e., keeping the number of permuted labels equal across subjects). This way, a random classifier was created for each subject. 2) For each subject, the classification score for the test set was then calculated using the random classifier from step 1. The classification outcomes of all subjects were averaged, and compared with the subject-averaged outcome based on the nonpermuted (veridical) classifier. Steps 1 and 2 were repeated 1000 times. A group-level *P*-value was calculated by counting how many times the random subject-averaged outcome was larger than the veridical subject-averaged outcome, and dividing this number by 1000.

To test whether the ambiguous color shifted toward a red or green color representation depending on object-color associations, the classifier was trained on either the 4 color runs or on the 4 typical object runs. The 4 runs belonging to the nontrained condition were tested as such that a correct classification corresponded to the typical-red condition being classified as the color red or the color red being classified as the typical-red condition (depending on whether training was performed on the color runs or on the typical object runs). Data for the 2 iterations (color to typical objects and typical objects to color) were averaged to obtain a more reliable estimate (i.e., smaller variance) of the effect. Again, data were tested using a permutation test as described above. To verify that the effect was due to typical object-color associations and not to other coincidental similarities between testing and training runs (e.g., order of presentation), we used the same procedure on the nonsense object condition. To obtain a classification score for the nonsense objects, one set of stimuli was arbitrarily chosen to represent the association with red (set A) and the other set was chosen to represent the association with green (set B).

To investigate whether other brain areas besides functionally defined early visual cortex are involved in object-color associations, we used a searchlight procedure to scan the entire brain using the same MVPA technique (Kriegeskorte et al. 2006). In this procedure, MVPA is performed on a spherical kernel with a radius of 4 voxels (257 voxels were weighted equally) throughout the entire brain (excluding the cerebellum and brainstem). To test each location in the brain, each voxel once served as the center of a spherical kernel. Classification was performed in each kernel separately, yielding a mean classification score (proportion correct) for each voxel in the brain for each subject. Again, the classifier was trained on one condition (color or typical/nonsense objects) and tested on the other condition, and the outcomes of

these 2 iterations were averaged. Because individual brain data were normalized to Talairach space, we were able to test the mean classification performance for each voxel across subjects. Mean classification performance in each voxel was tested against chance (0.50) using two-sided paired *t*-tests. First, single voxels were thresholded at a *P*-value of 0.01 (corresponding to a *t*-value of 3.25). Then, we determined whether a cluster of adjacent voxels – consisting of voxels touching in at least one corner in 3D space – exceeded our cluster threshold. To apply a cluster threshold, we used a boxplot (Frigge et al. 1989) to identify clusters that deviated in size from our sample based on the interquartile range of all cluster sizes (see Fig. 3A). We labeled these clusters significant, as their size exceeded the expected size based on the distribution of our sample.

## Results

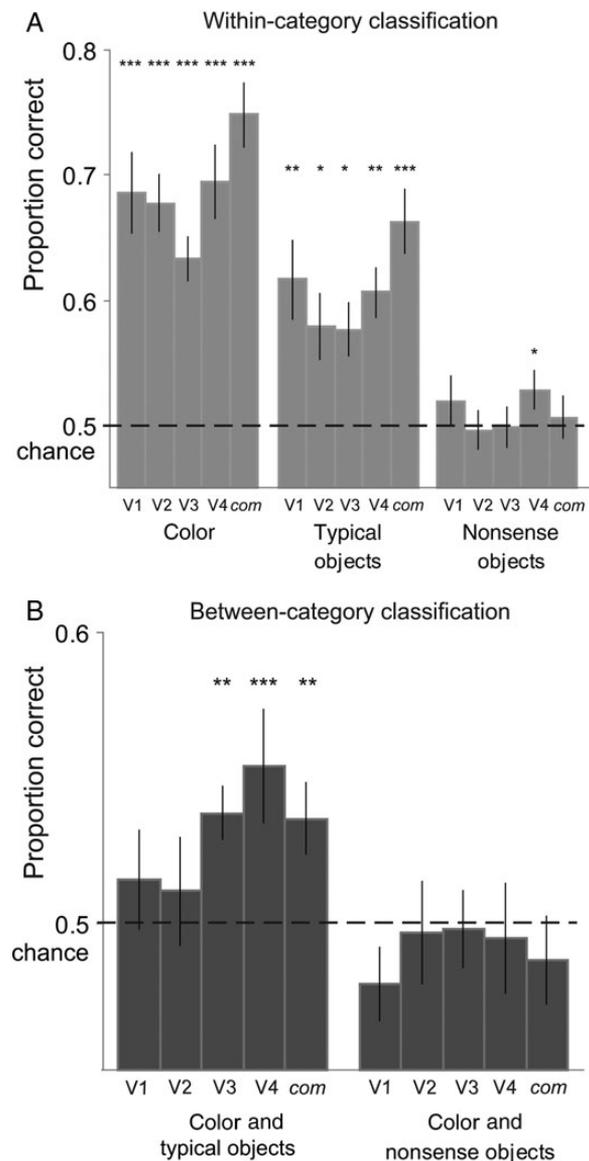
### ROI Decoding Results

To validate the MVPA technique with our particular dataset, we first determined whether we were able to correctly predict the presentation of color (red vs. green), typical objects (typical-green vs. typical-red) and nonsense objects (set A vs. set B) by themselves. Classification performance averaged over subjects is shown in Figure 2A. Red and green could be predicted for all 4 ROIs and when combining these ROIs (com), classification performance increased (all  $P < 0.001$ ). The 2 sets of real objects could be predicted from all visual areas as well (all  $P < 0.05$ ). The 2 types of nonsense objects could only be classified in V4 (V4:  $P = 0.048$ , V1–V3 and com:  $P > 0.05$ ). This indicates that there was not enough difference in spatial information to distinguish the 2 nonsense object sets (that were identical except for a rotation of each object) in V1–V3, while the difference in overall shape was processed by V4, albeit to a weak extent. Color categorization was superior to object categorization, in line with the idea that these low-level areas represent features rather than complete objects (Brewer et al. 2005; Wandell and Winawer 2011).

To investigate whether the representation of the ambiguous color shifted toward either red or green in the different object sets, we performed 2 between-category classifications. In these classifications, the classifier is trained on all the runs of one category (e.g., color) and tested on all the runs of the other category (e.g., objects). Data for the 2 iterations averaged (color to typical objects and typical objects to color) are shown in Figure 2B. Classification of the ambiguously colored typical objects as their typical color was significantly above chance for area V3 ( $P = 0.005$ ), V4 ( $P < 0.001$ ), and for the 4 ROIs combined ( $P = 0.005$ ). This suggests that, for the typical objects, the representation of the ambiguous color in V3 and V4 shifted toward either red or green, depending on the color associated with the specific typical object set.

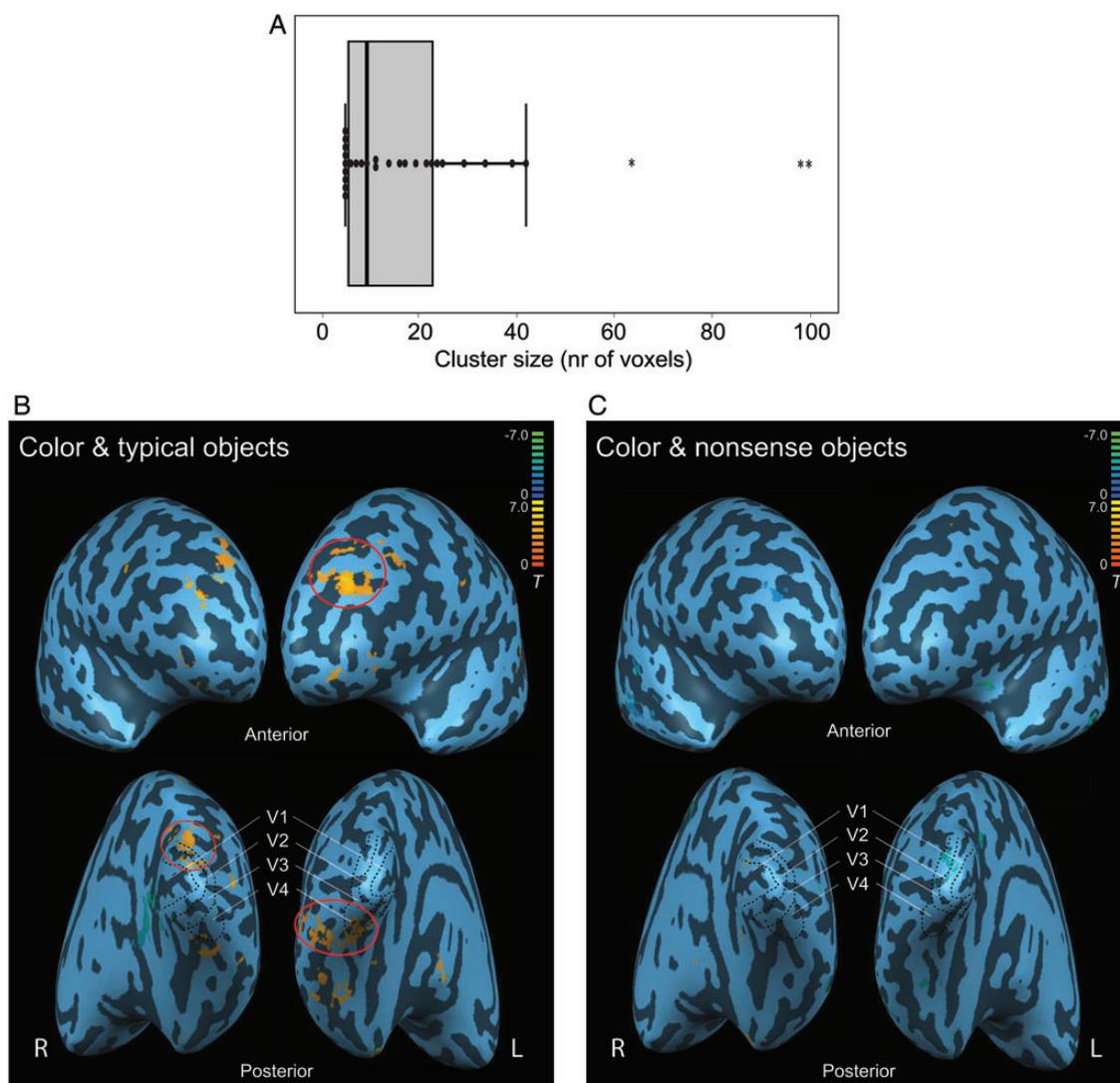
To verify that the effect was due to typical object-color associations, we used the same procedure on the nonsense object condition. To obtain a classification score for the nonsense objects, one set of stimuli was arbitrarily chosen to represent the association with red (set A) and the other set was chosen to represent the association with green (set B). In contrast to the typical objects, the classification between ambiguously colored nonsense objects and color was at chance performance (all  $P > 0.05$ ). As the ambiguous color was chosen to lie midway between red and green in CIE  $L^*a^*b$  color space, it was equally often classified as either red or green in both sets.

Together, these findings show that the representation underlying the ambiguous color in itself did not resemble



**Figure 2.** ROI analyses. (A) Within-category classification using a leave-1-run-out procedure for red and green (color), typical-red and typical-green objects (typical objects) and nonsense objects set A and B (nonsense objects) averaged over subjects. Classification of color and typical objects was significantly above chance (dotted line) for each ROI (V1–V4) and for all ROIs combined (com). Classification of nonsense objects only significantly exceeded chance in V4. (B) Classification between color and typical objects and between color and nonsense objects averaged over subjects and over training sets (see text). The between-category classification for color and typical objects was significantly above chance in V3, V4, and when all ROIs were combined (com). The between-category classification for color and nonsense objects did not deviate from chance. Error bars denote within-group standard errors. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

either red or green, but object knowledge influenced its neural representation in V3 and V4 such that it shifted toward the expected color. However, classification performance could also be driven by univariate differences, that is, differences in mean activity between ROIs. If, for example, the mean activity for the typical-red objects was higher than the mean activity for the typical-green objects, and at the same time the mean activity for the red geometrical shapes was higher than for the green geometrical shapes, this might drive the between-category classification results. To test this, we performed a



**Figure 3.** Searchlight analyses. (A) Cluster correction method used for the between-category (color and typical objects, and color and nonsense objects together) searchlight analyses. As the data were not normally distributed, we used a boxplot to identify clusters that deviated from the sample (using only clusters that had a voxelsize  $>5$ ). The advantage of using this method is that a boxplot is nonparametric, and thus does not make any assumptions about the underlying statistical distribution. The box indicates the interquartile range (IQR; left side = 1st quartile, right side = 3rd quartile), and the whiskers show the lowest and highest datum still in the  $1.5 \times \text{IQR}$  range. The thick line indicates the mean of the data. Three clusters exceeded the cluster threshold (stars), and these 3 clusters belonged to the color and typical objects condition. (B) The 3 significant clusters for the typical objects and color classification (circled in red, colors depict voxel  $t$ -values) were located in the right dorsal visual region, left ventral visual region, and left medial prefrontal region. (C) No clusters for the color and nonsense objects condition exceeded the cluster threshold.

univariate analysis, in which we tested the difference in activation between red and green, and between typical-red and typical-green for each ROI, using a general linear model. We found that there were no significant differences between red and green or the typical-red and typical-green object sets (Supplementary Fig. 2). Also, within participants, there was no relation between the direction of difference between red and green and the typical-red and typical-green objects (Supplementary Table 1). This suggests that the between-category classification performance for veridical red and green with ambiguously colored typical objects was not based on an incidental overlapping difference in mean activity between red and green, and the typical-red and typical-green object set. Thus, the correct classification of ambiguously colored typical objects as red or green depended on a shift in the specific voxel pattern response for the ambiguous color toward the voxel pattern response for red or green.

### Searchlight Decoding Results

To investigate which regions besides our functionally mapped lower visual regions were involved in the shift in color perception, we performed a searchlight analysis on both the typical objects and nonsense objects. We found 3 clusters for which the classification between colors and typical objects was significantly above chance (Fig. 3B): one right-lateralized visual dorsal cluster, one left-lateralized visual ventral cluster, and one left-lateralized prefrontal cluster. The visual dorsal cluster overlapped with area V2 and V3, and the visual ventral cluster overlapped with V4, thereby confirming the findings on V3 and V4 of our ROI analyses. Possibly, the patch of voxels that was involved in V2 was not large enough to yield a correct classification when all voxels in V2 were taken into the analysis. In addition to significant classification in our predefined ROIs, we found that the ventral cluster covering V4 also covered voxels anterior to V4, probably encompassing VO1

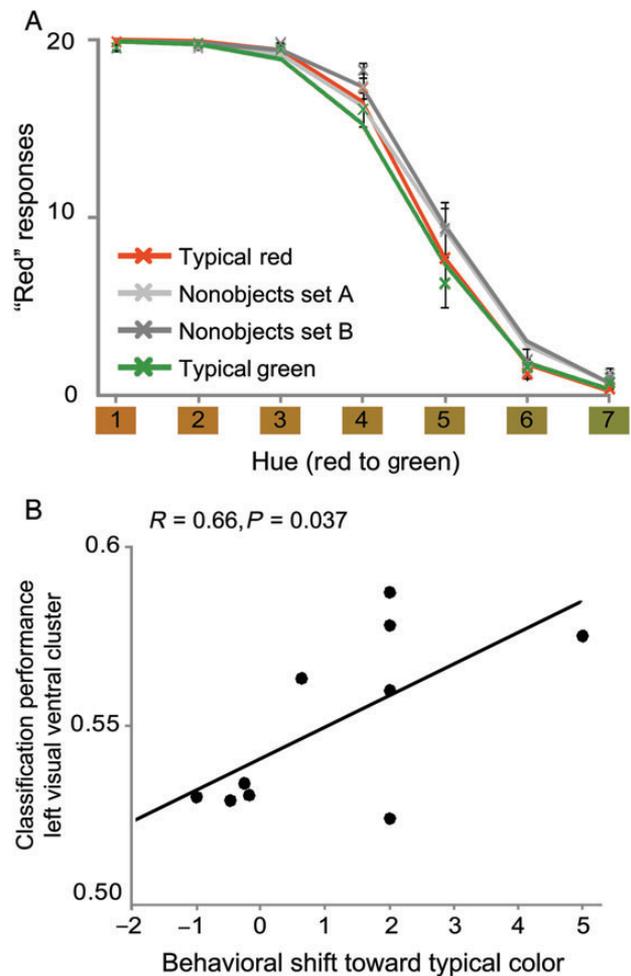
(Wandell and Winawer 2011), which is known to be involved in perceptual color processing just as V4 (Brewer et al. 2005; Bartels and Zeki 2008; Brouwer and Heeger 2009), and voxels in the left lateral occipital cortex (LOC), which is involved in object processing (Malach et al. 2002; Grill-Spector 2003). The visual dorsal region also encompassed a region anterior to V3. Post hoc, we tested whether this region might be V3A (Wandell and Winawer 2011) by defining V3A using the retinotopy and eccentricity mappings, but this was not the case (proportion correctly classified: 0.50,  $P > 0.05$ , results not shown). In addition to the involvement of visual areas, we found significant classification in left-superior and middle-frontal areas in the vicinity of the dorsolateral prefrontal cortex (DLPFC). No clusters survived the threshold for the nonsense object sets, confirming that these results are specific to the experimental manipulation of object category (see Fig. 3C).

### Behavioral Results

To determine whether the typical objects evoked a behavioral change in perception that coincided with the neural findings, after scanning subjects performed a task in which they had to indicate whether they perceived a color (from a 7-scale continuum) presented on an object as either red or green. Figure 4A shows the number of “red” responses for each of the 7 hues on the 4 object sets, together with a fit of the data based on a binomial regression (logit transformed to obtain continuous data). For these 10 subjects, there was a trend towards giving more “red” responses for the typical-red objects than for the typical-green objects when presented in hue 4, which was the color that was used in the fMRI experiment ( $t_{(9)} = 2.1$ ,  $P = 0.067$ ). The 2 nonsense object sets were labeled “red” more often than both the typical-green and the typical-red objects. This suggests that the perception of color on these objects leans toward red, possibly because there is less contrast in the nonsense objects than in the typical objects. We then investigated whether the behavioral effect on the color that was presented in the fMRI experiment correlated with our neural classification measure in one of the significant ROIs or clusters from our MVPA. We found that classification scores for the left visual ventral cluster (encompassing V4/VO1/LO) correlated positively with the behavioral effect (Pearson’s  $R = 0.66$ ,  $P = 0.037$ ; Fig. 4B), supporting the hypothesis that these areas are involved in subjective color perception by combining memory for colors with incoming color information (Shapley and Hawken 2011). No significant correlations were found for the nonsense object set.

### Discussion

In this study, we investigated which brain regions are influenced by prior knowledge about objects, thereby shaping our subjective color experience (Hansen et al. 2006; Mitterer and de Ruyter 2008). Using MVPA, we were able to classify the response to an ambiguous color – lying midway between red and green – as red when presented on typical red objects, while that same color was classified as green when presented on typical green objects. In contrast, when the ambiguous color was presented on 2 sets of nonsense objects that did not have any color associations, the color could not be classified as either red or green. When using functionally defined ROIs in visual cortex, we found the areas to be involved in this transformation were located in lower level visual areas V3 and V4.



**Figure 4.** Behavioral results. (A) Number of “red” responses for each hue in each object set (total of 4 objects  $\times$  5 repetitions = 20 observations per stimulus condition). Crosses indicate the response averaged over all subjects. Solid lines represent the fit of the data using a binomial regression (logit transformed). Subjects gave more “red” responses to typical red objects than to typical green objects. The nonsense object sets (nonobjects set A and B) were labeled red more often than both the typical object sets. Hue 4 is the ambiguous color that was used in the fMRI experiment. Error bars denote within-group standard error for the unfitted data (crosses). (B) Correlation between the behavioral effect and classification performance for the left visual ventral cluster (V4/VO1/LO). The behavioral shift was calculated as the difference in “red” responses for hue 4. Note, due to the binomial fit on the logit transformed data, the number of observations is continuous.

This shows that subjective experience at least partly overrides the representation of physical stimulus properties at a relatively early stage of the visual processing hierarchy. When using a whole-brain searchlight approach, we confirmed the contribution of early visual cortex to object-color representations, but also found above chance classification performance in frontal regions (in the vicinity of the DLPFC). This suggests that subjective color experience might be mediated by object-color knowledge through involvement of frontal areas as well.

The neural correlates found in this study coincide with previous work showing that responses to color in V4 and VO1 (and somewhat in V3) progress through perceptual color space and not physical color space (Brouwer and Heeger 2009). In V1 and V2 on the other hand, responses to color can be easily decoded, but do not progress through color space in the same manner as in V4. This suggests that V4 represents perceptual color rather than physical color input, whereas responses in V1

and V2 might be mainly driven by physical color input. In the present study, we found that object information shifted the representation of a single color toward either red or green in V3 and V4, showing that the representations in these areas are not only dominated by perceptual color, but are modulated by object-color knowledge. In V1 and V2, on the other hand, veridical red and green representations were decodable, but the representation of the ambiguous color did not shift according to object-color associations. Possibly, representations in V1 and V2 are dominated by physical color input, and are not influenced by object information. This might also explain why combining all ROIs improved classification performance for within-category classification (red vs. green), but did not improve between-category classification (colors with objects); in the latter case, the representations in V1 and V2 diverges from that in V3 and V4 and combining the information represented in these areas does not lead to superior performance.

In addition to investigating functionally defined ROIs, we performed a searchlight analysis and found that the voxel pattern associated with the ambiguous color also shifted toward red or green in left VO1 and left LOC. VO1 (or V4 $\alpha$ /V8) is an area that is known to be involved in color processing as well and has been suggested to be involved in perceptual color representations just as V4 is (Brewer et al. 2005; Bartels and Zeki 2008; Brouwer and Heeger 2009). The LOC sits higher up in the visual hierarchy and is known to be involved in object processing (Malach et al. 2002; Grill-Spector 2003). LOC might be involved in the shift in color perception because the perception of the typical objects is coupled to a specific color, thereby automatically activating the associated object-color representation. Moreover, individual classification scores in this cluster correlated with the behavioral color naming effect, supporting the claim that these areas are involved in subjective color experience by combining incoming color information with existing object-color memories (Shapley and Hawken 2011).

In addition to visual cortex, our searchlight analysis revealed significant between-category classification in prefrontal areas including the DLPFC. There are different possible explanations for this finding: one possibility is that visual information is processed by prefrontal cortex in the same manner as it is in visual areas, and any shift in perception will therefore be reflected in these regions as well. However, this is unlikely, since this would be a redundant operation and there is not much evidence showing that frontal cortex is involved in primary analysis of visual information (but see Bar et al. 2006). Alternatively, prefrontal cortex has been shown to be involved in retrieval of associative long-term memory items (Hasegawa et al. 1998; Ranganath et al. 2004). It could be that when viewing a certain stimulus, associated concepts are activated and therefore, when viewing an ambiguously colored object, the memory for the typical color of that object is activated as well. The process of perceiving colors differently depending on their semantic context might be facilitated by predictive signaling from these higher level brain areas toward lower level brain regions (Rao and Ballard 1999).

As a control, we showed that nonsense objects could not be classified using the neural representations for red and green. However, the within-category classification for nonsense objects was only marginally significant in V4, and not significant in V1–V3. Because the within-category classification did not produce a large effect, it only weakly substantiates the absence of a type I error on the within-category classification

results for typical objects. Therefore, it also carries less weight as a control for the between-category classification results for the typical objects. As a result, the outcome of the between-category control classification is less informative, and would have given stronger evidence if the within-category classification had been successful, or more significant, in all visual areas. However, the fact that within-category and between-category classifications worked for the typical object and color conditions but not for the control condition, suggests that neither within-category nor between-category classifications for color or typical objects is explained by unforeseen confounds in the classification procedure that are unrelated to the experimental manipulations (e.g., an unforeseen impact of stimulus ordering, phasic changes in attention, or classifier bias).

The behavioral effect that we found in this study was statistically weak. Others investigating a similar shift in ambiguous color perception have found stronger results (Mitterer and de Ruiter 2008). A reason for this might be that red and green are perceptually quite distinct, and thus even though physically, an ambiguous color might lie exactly between red and green, it is hard to define an ambiguous color that has the same perceptual distance from both red and green. This can be seen by the fact that the curve for the amount of given red responses is quite steep (Fig. 4A): there seems to be an abrupt shift in labeling a color as either red or green. However, we specifically chose to contrast red and green, because these are neurally more distinct than, for example, orange and yellow. We therefore predicted that if we find a shift in neural color representation, it would be most evident along the red–green continuum where the inherent difference is large. In that sense, it might not be surprising that we find neural effects that are accompanied by somewhat weaker behavioral effects.

Another reason for the small behavioral effect might be that line drawings produce a smaller memory color effect than for example grayscale photographs (Witzel et al. 2011). In this study, we specifically chose to use line drawings and not photographs, because we wanted to minimize effects due to local contrast differences. For grayscale photographs, overall luminance can be equalized between stimuli, but local contrast differences remain. This might specifically affect neural operations in lower level areas involved in processing contrast. Recently, Bannert and Bartels (2013) conducted an experiment similar to the current study in which they used grayscale photographs. Indeed, they were able to decode photographs using color representations in V1, but not in V4. This difference in results might be explained by the fact that local contrast differences were processed and interpreted specifically in V1, which resulted in color (or luminance) information that was congruent with the actual colors. Another possibility is that in our study, the representation of the ambiguous color could be overruled and altered in V4, but not in V1, while with grayscale photographs, only V1 is affected.

The present study shows that object-color associations influence color processing in visual areas representing colors and objects, as well as in frontal areas associated with memory. The fact that representations in early visual areas are modified by object knowledge suggests that subjects not only categorize a color according to semantic expectations, but actually “perceive” a color differently depending on the object it is presented on. Moreover, this effect occurs instantly, as the same ambiguous color can be represented as green or red within

one experiment. Such a process might be supported by predictive signaling from higher level to lower level brain areas (Rao and Ballard 1999). The current results reveal that the brain shapes our subjective experience by rapidly incorporating world knowledge, and altering neural responses in the cortical areas that are involved in the initial stages of visual processing.

### Supplementary Material

Supplementary Material can be found at <http://www.cercor.oxfordjournals.org/online>.

### Funding

This project was funded by an Advanced Investigator Grant (ERC Investigator Grant DEFCON1 230355) from the European Research Council to V.A.F.L.

### Notes

The authors declare no competing financial interests. We thank Jasper Wijnen for programming the experimental task. We thank Gijs Brouwer for his helpful comments on data analyses. *Conflict of Interest*: None declared.

### References

- Bannert MM, Bartels A. 2013. Decoding the yellow of a gray banana. *Curr Biol*. 23:2268–2272.
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmidt AM, Dale AM, Halmalainen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E. 2006. Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA*. 103:449–454.
- Bartels A, Zeki S. 2008. The architecture of the colour centre in the human visual brain: new results and a review. *Eur J Neurosci*. 12:172–193.
- 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.
- Brouwer GJ, Heeger DJ. 2009. Decoding and reconstructing color from responses in human visual cortex. *J Neurosci*. 29:3992–14003.
- Frigge M, Hoaglin DC, Iglewicz B. 1989. Some implementations of the boxplot. *Am Stat*. 43:50–54.
- Gegenfurtner KR, Kiper DC. 2003. Color vision. *Neuroscience*. 26:181–206.
- Goebel R, Esposito F, Formisano E. 2006. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp*. 27:392–401.
- Grill-Spector K. 2003. The neural basis of object perception. *Curr Opin Neurobiol*. 13:159–166.
- Hansen T, Olkkonen M, Walter S, Gegenfurtner KR. 2006. Memory modulates color appearance. *Nat Neurosci*. 9:1367–1368.
- Hasegawa I, Fukushima T, Ihara T, Miyashita Y. 1998. Callosal window between prefrontal cortices: cognitive interaction to retrieve long-term memory. *Science*. 281:814–818.
- Heywood CA, Kentridge RW. 2003. Achromatopsia, color vision and cortex. *Neurol Clin N Am*. 21:483–500.
- Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. *Proc Natl Acad Sci USA*. 103:3863–3868.
- Malach R, Levy I, Hasson U. 2002. The topography of high-order human object areas. *Trends Cogn Sci*. 6:176–184.
- Mitterer H, De Ruyter JP. 2008. Recalibrating color categories using world knowledge. *Psychol Sci*. 9:629–634.
- Parkes LM, Marsman JBC, Oxley DC, Goulermas JY, Wuerger SM. 2009. Multivoxel fMRI analysis of color tuning in human primary visual cortex. *J Vis*. 9(1). doi: 10.1167/9.1.1.
- Poldrack RA, Mumford JA, Nichols TE. 2011. Handbook of functional MRI data analysis. Cambridge, UK: Cambridge University Press.
- Ranganath C, Cohen MX, Dam C, D'Esposito M. 2004. Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J Neurosci*. 24:3917–3925.
- Rao RP, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*. 2:79–87.
- Seymour K, Clifford CWG, Logothetis NK, Bartels A. 2009. The coding of color, motion, and their conjunction in the human visual cortex. *Curr Biol*. 19:177–183.
- Shapley R, Hawken MJ. 2011. Color in the cortex: single- and double-opponent cells. *Vision Res*. 51:701–717.
- Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Wandell BA, Winawer J. 2011. Imaging retinotopic maps in the human brain. *Vision Res*. 51:718–737.
- Witzel C, Valkova H, Hansen T, Gegenfurtner KR. 2011. Object knowledge modulates colour appearance. *Iperception*. 2(1):13.
- 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.
- Zeki S, Marini L. 1998. Three cortical stages of colour processing in the human brain. *Brain*. 121:1669–1685.