Temporal dynamics of the neural mechanisms for encoding hue and luminance contrast uncovered by magnetoencephalography

Katherine Hermann[1*^], Isabelle Rosenthal[1*#], Shridhar Singh[1*], Dimitrios Pantazis[2], Bevil R. Conway[1,3]

1. Laboratory of sensorimotor research, National Eye Institute, Bethesda MD 20892
2. McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge MA 02139
3. National Institute of Mental Health, Bethesda MD 20892

*equal contribution

^current address, Stanford University

#current address, Caltech graduate program in Computation and Neural Systems

Author contributions: KH and BRC conceived and designed the experiments; KH and IR collected the data; KH, IR, SS analyzed the data; IR, SS and BRC made the figures; DP provided expertise and resources for MEG; BRC wrote the paper.

Hue and luminance contrast are the most basic visual computations, and are reflected in the earliest layers of convolutional neural networks, yet the extent to which they are extracted by the same or separate circuits in the brain, and the timing of these neural computations, is unknown. Here we answer these questions using multivariate analyses of human brain responses measured with magnetoencephalography. We report three discoveries. First, hue and luminance contrast could be decoded independently, indicating these computations are somewhat separable. Hue was computed about 15-24ms after luminance contrast. Second, representations of hue showed relatively greater generalization across time and were more sustained, providing the first neural correlate of the perceptual preeminence of hue over luminance contrast in grouping objects. Finally, luminance contrast could be decoded less well for hues associated with daylight (orange and blue), suggesting that color-constancy mechanisms are adapted to natural lighting.

The most basic computations performed by the visual system yield hue and luminance contrast (Ennis and Zaidi, 2019, Smet et al. , 2016). These computations are reflected as separate filters in early layers of convolutional neural networks (Flachot and Gegenfurtner, 2018, Krizhevsky et al. , 2012, Rafegas and Vanrell, 2017, Zeiler and Fergus, 2014). *Hue* is that property of color referred to by a color name (e.g., blue, green). *Luminance contrast* is the gray-value equivalent—how light or dark a color is—which is, importantly, distinct from absolute luminance. Absolute luminance derives from the spectral sensitivity of L and M cones; luminance contrast depends on post-receptoral computations (Kuffler, 1953). A long-standing question is whether the neural computations that support hue and luminance contrast are independent (Clifford et al. , 2003, Favreau and Cavanagh, 1981, Goddard et al. , 2019); another unanswered question is how long the brain takes to compute hue and luminance contrast (Maunsell et al. , 1999).

One prominent theory, succinctly articulated by Gegenfurtner (2003), is that hue is "processed not in isolation, but together with information about luminance [contrast] and visual form, by the same neural circuits, to achieve a unitary and robust representation of the visual world" (Gegenfurtner, 2003). This position is supported by neurophysiological recordings in the lateral geniculate nucleus (LGN) of monkeys. Parvocellular LGN cells multiplex tuning for luminance contrast and color opponency (Reid and Shapley, 2002, Wiesel and Hubel, 1966). Indeed, the response properties of parvocellular cells correspond to the relative spatial frequency of luminance-contrast vision (high-pass) and color vision (low-pass) (De Valois and Switkes, 1983, Granger and Heurtley, 1973, Mullen, 1985, van der Horst and Bouman, 1969). But besides parvocellular neurons, the LGN contains magnocellular cells, which respond to luminance contrast and are not cone opponent; and koniocellular neurons, some of which respond to S-cone signals and may represent a distinct chromatic channel (Lee et al. , 1989, Martin et al. , 1997). These observations animate an alternative idea, that the LGN encodes hue and luminance contrast in separable channels (Dobkins, 2000).

The LGN data are inconclusive not only about the extent to which luminance contrast and hue are encoded by separate channels, but also about the amount of time that these computations take. On the one hand, if parvocellular neurons are the basis for computing both hue and luminance contrast, these computations should take the same amount of time. On the other hand, if color and luminance contrast are encoded by different LGN channels, one might expect luminance contrast to be computed earlier than hue because

magnocellular neurons have shorter latencies than parvocellular neurons. But because there are relatively fewer magnocellular neurons, their latency advantage may be lost through convergence in visual cortex (Maunsell et al., 1999). Clues to the neural mechanisms of color and luminance contrast are provided by univariate visual evoked potential measurements to equiluminant and achromatic stimuli (Rabin et al. , 1994). But it has not been possible to infer from these experiments the underlying neural mechanisms because both main subcortical channels respond to equiluminant stimuli (Logothetis et al. , 1990). Moreover, such experiments are inconclusive about timing because response latency depends on stimulus contrast, and there is no accepted metric for equating color contrast and luminance contrast (Shevell and Kingdom, 2008).

Neurophysiological data from visual cortex, the next stage of visual processing downstream of the LGN, have not resolved the central questions and have uncovered a curious paradox. The great majority of V1 cells discriminate precisely the orientation or direction of movement of a stimulus but have no marked hue selectivity (Conway, 2001, Horwitz and Hass, 2012, Hubel and Wiesel, 1968, Johnson et al. , 2004, Lennie et al. , 1990, Nealey and Maunsell, 1994, Solomon and Lennie, 2005). The V1 data imply that luminance contrast is the overwhelmingly dominant feature encoded by visual cortex, yet behaviorally, hue trumps luminance contrast under many situations. A preeminence of hue is evident in the difficulty one encounters when trying to match the brightness of two different hues (De Valois and Switkes, 1983). People with normal color vision will typically group stimuli by hue rather than luminance contrast. Consider the eight spirals in **Figure 1a**; most people group them by rows not columns. There is presently no neurophysiological correlate for the preeminence of hue over luminance contrast. Most V1 cells appear to receive a mixture of inputs from subcortical channels (Nealey and Maunsell, 1994), such that most cells appear to multiplex luminance contrast and color (Gegenfurtner, 2003, Johnson et al., 2004, Thorell et al. , 1984). Nonetheless, a minority of V1 neurons retain clear cone-opponent spatial structure with striking color tuning (Conway, 2001, Lafer-Sousa et al. , 2012), consistent with the idea that hue and luminance contrast could be encoded by parallel pathways through V1 (Nassi and Callaway, 2009).

Although standard representations of color depict hue and luminance contrast as independent dimensions, there are situations in which one might expect these dimensions to interact. Natural illuminants, which range from orange to blue chromaticities, confound chromaticity and luminance: sunlight, associated with orange chromaticities, is associated with high luminance contrast. Sunlight casts shadows that reflect blue

skylight, so blue is associated with low luminance contrast. This confound may underlie the psychological association of warm colors with “light” and cool colors with “dark” (Lindsey and Brown, 2006), and raises a question: does the brain less reliably compute luminance contrast from hues associated with daylight (orange/blue) compared to anti-daylight (green/pink), as one might expect if the brain is adapted to natural lighting conditions? Addressing this question has important implications for understanding color constancy (Delahunt and Brainard, 2004, Lafer-Sousa et al. , 2015, Lafer-Sousa et al., 2012, Pearce et al. , 2014, Winkler et al. , 2015). Our goal in the present work was to flip the traditional logic: rather than using psychophysics to infer neural mechanisms, we aimed to directly measure neural responses using magnetoencephalography (MEG) in humans, coupled with multivariate analysis (Carlson et al. , 2013, Cichy et al. , 2014, Isik et al. , 2014, Sandhaeger et al. , 2019, van de Nieuwenhuijzen et al. , 2013, Wardle et al. , 2016), to address fundamental questions about perception.

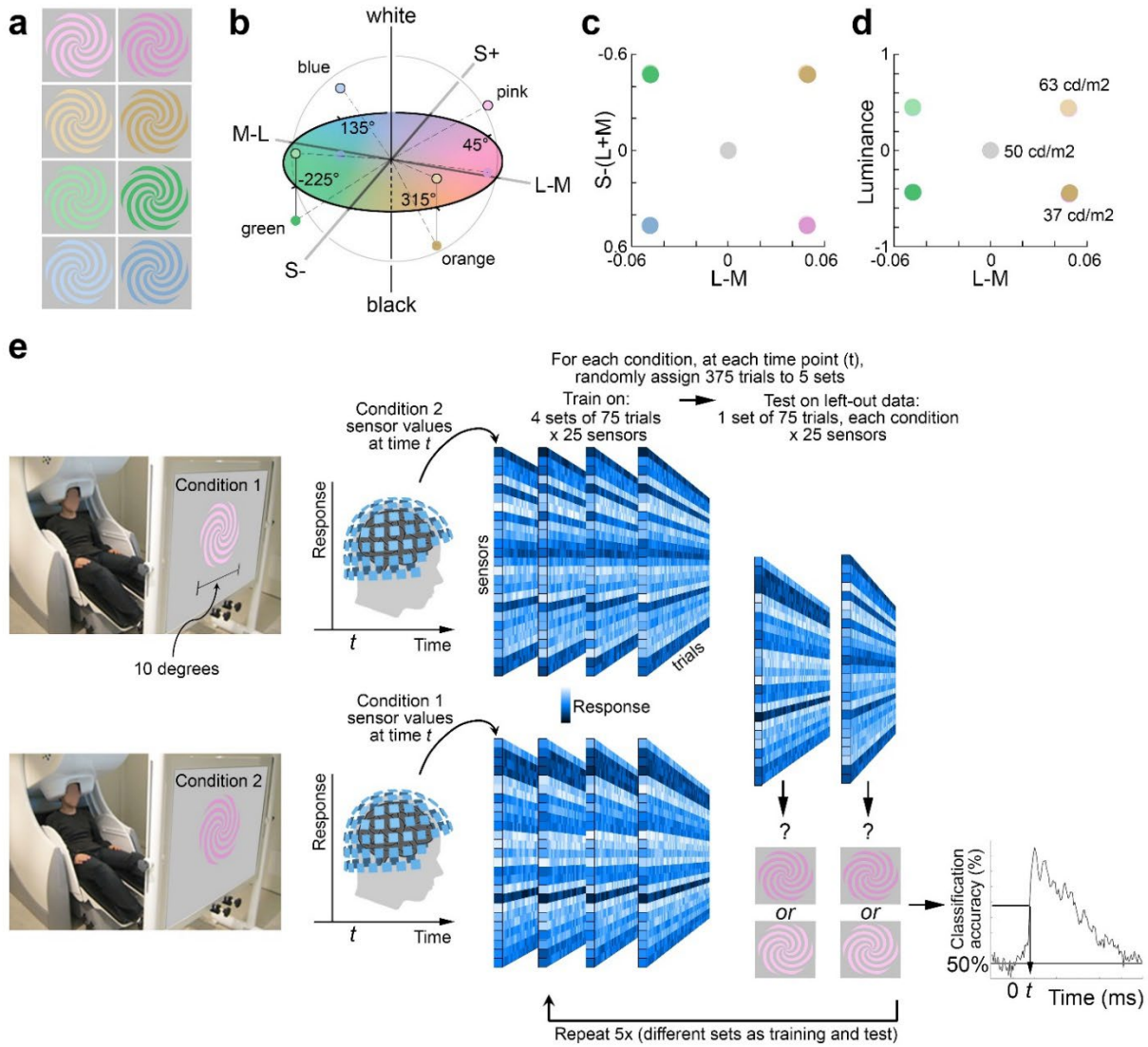

Figure 1. Stimulus specification and experimental paradigm

a) The stimuli consisted of 8 colored spirals, four hues (rows) at two luminance levels (columns). **b)** Cone-opponent (DKL) color space showing the location of the eight colors used in the experiments. The axes of the space are defined by the cardinal chromatic mechanisms that reflect the color properties of the major categories of neurons in the retina. The four hues are the intermediate hues: the magnitude of modulation along the two cardinal mechanisms is the same for all stimuli. **c)** Cone-contrast of stimuli used, plotted in the equiluminant plane of DKL color space; values are cone-contrast computed relative to adapting background (gray, at 0,0); note the light stimuli are plotted behind the dark stimuli. **d)** Contrast of stimuli used, plotted in the luminance plane DKL color space; luminance values are normalized to the maximum contrast of the monitor (absolute luminance values given in cd/m^2); note the S-increment stimuli are plotted behind the S-decrement stimuli. **e)** Schematic showing the experimental paradigm, including a participant in the MEG scanner with one of the eight stimuli on the screen, and simulated data (for illustration purposes). Each stimulus (i.e. "condition") was presented for 116ms with 1 second between presentations. Stimuli were pseudo-randomly interleaved, with 500 presentations of each stimulus over two recording sessions. Trials during eye blinks or other artifacts were removed, and the remaining trials were randomly subsampled to yield 375 trials per condition. Sensor data were averaged into 5ms bins within a time window of 200ms before stimulus onset to 600ms after stimulus onset. For the analysis, at each time point (*t*) in the 800ms time window, the 375 trials were divided into 5 sets of 75. Four sets were used to train the classifier, and 1 set was used to test the classifier. The procedure was repeated for the 5 cross-validation splits; and the entire procedure was repeated 50x with different random assignments of the 375 trials into the 5 sets.

Results

The experiments were designed to enable a decoding analysis of MEG responses, to answer the over-arching questions: (1) how much time does the brain take to compute luminance contrast and hue? And (2) can the luminance contrast and hue of a stimulus be decoded independently given the pattern of MEG activity elicited by the stimulus? We measured MEG responses in 18 participants while they were shown brief presentations of 8 colored spirals (**Figure 1a**). The stimuli were defined by cone-opponent dimensions that reflect how the retina encodes chromatic information (**Figure 1b-d**) (Derrington et al. , 1984, MacLeod and Boynton, 1979). The 8 colors consisted of four hues at two luminance levels (4 X 2 design). The luminance contrast of all the stimuli was the same (26%) but varied in sign (light or dark) relative to the neutral adapting background. If hue and luminance contrast are encoded by separable neural mechanisms, it should be possible to decode hue even if the MEG data used to train the classifiers were elicited by stimuli that differed in luminance contrast from the test stimuli; and it should be possible to decode luminance contrast even if the classifiers were trained using data elicited by stimuli that differed in hue from the test data. The time course should tell us about the relative stage in the visual-processing hierarchy at which hue and luminance contrast representations are encoded and/or the relative amount of recurrent processing required for each computation. Alternatively, if hue and luminance contrast are encoded together, it should be possible to decode specific hue-luminance combinations, but not each dimension separately.

We used maximum correlation coefficient classifiers implemented in the Neural Decoding Toolbox (Meyers, 2013) (see Methods). We trained separate classifiers for each participant, and separate classifiers at each time point relative to stimulus onset. All analyses were cross validated (see **Figure 1e**), yielding plots that show how the representations unfold over time. Participants were told to maintain fixation throughout stimulus presentation and to blink during designated times (**Figure 1e,** left). Data during eye blinks or breaks in fixation were removed (see Methods). To control fixation and attentional state, participants engaged in a 1-back hue-matching task: every 3-5 trials, the participants were queried with a "?" on the screen to report via button press whether the two preceding stimuli matched. Pilot experiments showed no difference in data obtained using a 1-back luminance-contrast matching task (**SI Figure 1**). We used a spiral-shaped stimulus to avoid cardinal or radial response biases (Brouwer and Heeger, 2009, Mannion et al. , 2009, Seymour et al. , 2010).

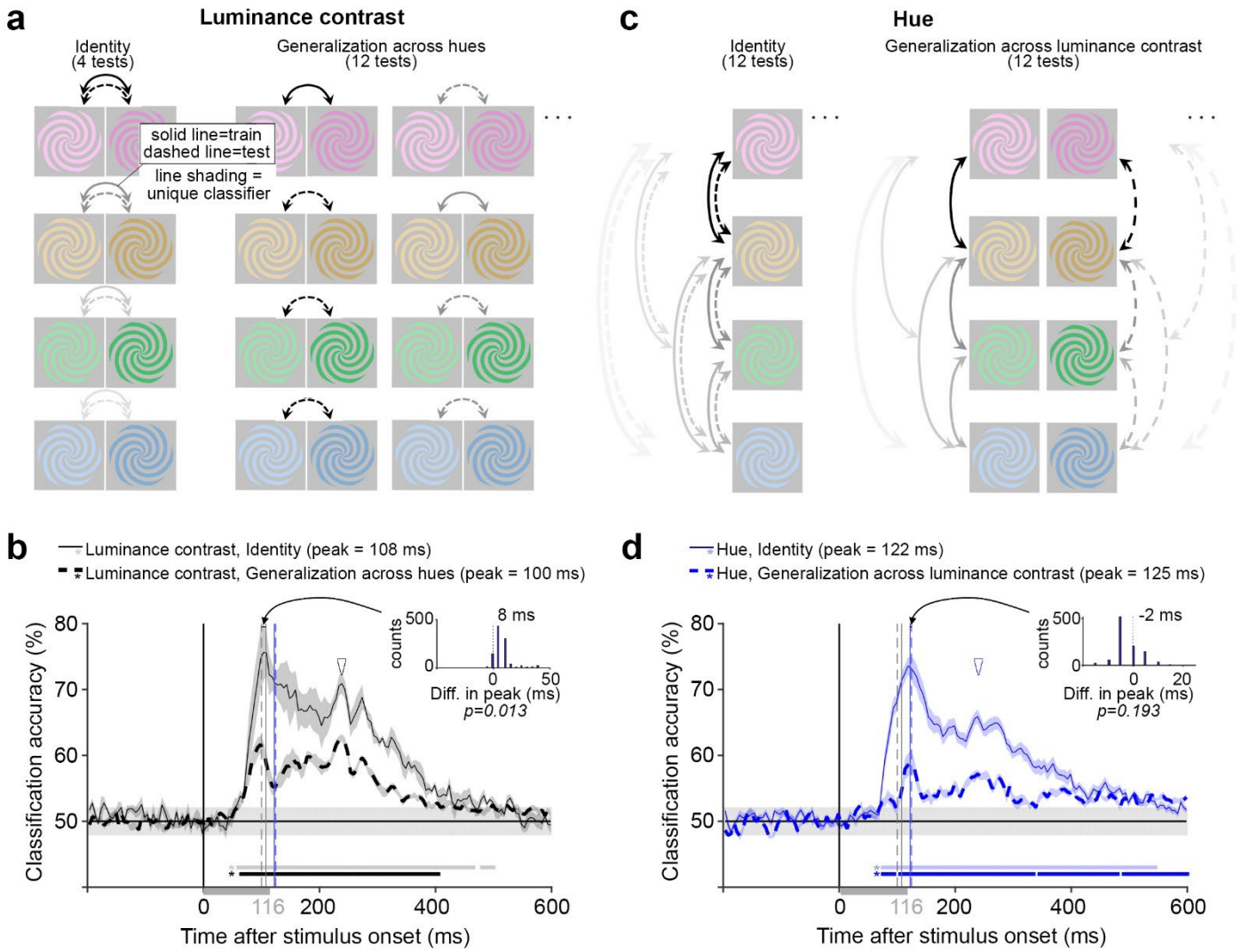

Figure 2. Decoding luminance contrast and hue from MEG data

a) Decoding luminance contrast. The schematic at the top illustrates the sets of analyses; the graph at the bottom shows the data. Participants were presented with 8 colored spirals: four hues (pink/orange/green/blue) at two luminance contrast levels (light/dark). Classifiers were trained to determine the extent to which the MEG response to a given color is informative of the luminance contrast carried by the same hue (4 identity problems) or by other hues (12 generalization problems). Each binary classifier was trained to distinguish whether a light or dark stimulus had been presented given patterns of MEG sensor activations. For the four identity problems, classifiers were trained and tested on the same hue, one classifier for each hue. For the generalization problems, twelve classifiers were trained and tested on different hues, one for each permutation of hues into training and testing sets. In the graphic, the line thickness and shading of the arrows identifies a unique classifier. **b)** The average performance across the four identity problems (solid line) and the 12 generalization problems (dashed line). The individual problems were evaluated for each participant separately. For each problem, we averaged the decoding performance across participants. The traces in the graph were generated by averaging the 1000 bootstrapped samples of the four identity problems (solid line) and averaging the 1000 bootstrapped samples of the twelve generalization problems. Shading around the 50% chance line shows the 95% CI of the decoding performance prior to stimulus onset (baseline); the stimulus duration was 116 ms (gray bar along the x axis). The inset shows the difference in peak (identity minus generalization, mean = 8 ms) for the 1000 bootstrapped comparisons. The identity problem had a slightly longer time-to-peak (p=0.013). Peak of the identity

problem, solid vertical gray line, 108ms [100 140]; first peak of the generalization problem, dashed gray line, 100ms [95 105]. Open arrowhead shows the peak in decoding corresponding to cessation of the stimulus. The horizontal lines demarcated by asterisks show the time points at which decoding was above chance. **c)** Decoding hue. Format as in panel a. At each level of luminance contrast (e.g. dark), a binary classifier was trained to determine which of two hues (e.g. pink or orange) had been presented. The classifiers were then tested on held-out trials in which the luminance contrast (e.g. dark) was the same as at train time (12 identity problems) or in which it was different (e.g. light), requiring generalization of hue across luminance contrast (12 generalization problems).**d)** The average performance across the 12 sets of identity problems (solid line) and 12 sets of generalization problems (dashed line). The inset shows the difference in peak (identity minus generalization, mean = -2 ms, p=0.193) for the 1000 bootstrapped comparisons. The time to peak was about the same for both identity and generalization (dashed vertical blue line, 125ms [115 130]; solid vertical blue line, 122ms [110 130]; not different, p=0.2). Other conventions as for panel (b).

The eight colors ("conditions" in **Figure 1e**) were presented in pseudo-random order for 116ms with 1 second of the gray background between presentations. We collected responses to a very large number of trials of each condition (N=500), removed trials with artifacts such as eye blinks, and randomly subsampled the remaining trials to obtain 375 trials per condition. **Figure 1e** is a cartoon illustrating an analysis in which classifiers were trained to decode luminance contrast given patterns of MEG data elicited by bright and dark pink; the classifiers were tested on separate data elicited by the same stimuli, bright and dark pink. The results reveal the classification accuracy for luminance carried by a specific hue (pink). We refer to this as a luminance-contrast identity problem because the hue of the stimuli from trials used to train the classifier is identical to the hue of the stimuli in the test trials (**Figure 2a, left**). In other tests of luminance-contrast decoding, the hues of the stimuli differed between trials used to train versus test the classifier. For example, classifiers were trained using patterns of MEG activity elicited by light and dark pink but tested using activity elicited by light and dark blue, or light and dark orange, or light and dark green. We refer to these as luminance-contrast generalization-across-hue tests (or problems) since they uncover the extent to which luminance contrast can be decoded separately from hue (**Figure 2a, right**). In other analyses we determined the extent to which classifiers could decode hue identity (**Figure 2c, left**) and hue generalized across luminance-contrast (**Figure 2c, right**). All analyses involved binary classifiers to facilitate a direct comparison of the results among the different tests. The arrows in **Figure 2a, 2c** show the stimuli associated with the training of classifiers (solid arrows) and the stimuli associated with the corresponding tests (dashed arrows).

Decoding luminance contrast

The data enabled us to perform four luminance identity tests (**Figure 2a left**) and twelve luminance generalization-across-hue tests (**Figure 2a right**). **Figure 2b** shows the average bootstrapped performance across the identity tests (solid line) and the generalization tests (dashed line; shading around each trace shows the bootstrapped standard error). The shading around the 50% chance line shows the 95% CI of the decoding performance prior to stimulus onset (baseline); the stimulus duration was 116 ms. The horizontal black and gray lines indicated by asterisks show the time points at which decoding performance rose above chance, defined as when classification accuracy crossed the upper 95% CI bound of the background decoding performance for more than 5 consecutive bins. The results show that classification accuracy was significantly above chance for luminance contrast for both types of tests (the dashed line rises above the upper bound of the 95% CI of the baseline). Thus it was possible to decode luminance contrast independent of hue, which supports the hypothesis that the brain has a spatial representation of luminance contrast that is separate from the representation of hue. Peak decoding accuracy was higher for the luminance-contrast identity tests compared to the luminance-contrast generalization tests (76% [70 84] versus (62% [59 64]; square brackets contain the 95% CI). This result supports the hypothesis that the brain also has a representation of luminance contrast that is combined with the representation of hue.

Despite the lower peak accuracy for the generalization problem, the time to peak for this problem was slightly earlier than for the identity decoding problem (dashed vertical gray line, 100ms [95 105]; versus solid vertical gray line, 108ms [100 140]; p=0.013; confidence limits computed by 1000 bootstrap draws of the individual problems). The generalization problem showed a pronounced dip following the peak, which occurred at the same time as the peak in decoding for the hue generalization problem described in the next section. The time point of the peaks in the hue decoding problems are indicated with vertical solid and dashed blue lines in Figure 2c, to facilitate comparison of the hue and luminance-contrast decoding results.

Latency of the decoding problems was determined as the time point at which classification accuracy rose 2.5 standard deviations above background decoding for 5 consecutive time bins, bootstrapped 1000x over the subproblems. The latency was not different for the luminance-contrast identity problem (64ms) versus the luminance-contrast generalization problem (69 ms; p=0.26). The time points denoted by asterisks in Figure 2b are when the average bootstrapped classification curve crossed the upper 95% CI bound of

background decoding for 5 consecutive time bins. By that measure, the latency of the identity problem was 60 ms and the generalization problem was 65 ms.

The graph uncovers two other findings. First, decoding for the identity problem was significant for a slightly longer period than for the generalization problem (the gray horizontal line demarcated by an asterisk at the bottom of the graph is slightly longer than the black horizontal line; these lines show time points at which the decoding performance crossed the upper 95% CI of baseline decoding for five consecutive 5ms time bins). And second, both the identity and generalization problems had two prominent decoding peaks: an initial peak (curved arrow) and a second peak (open arrowhead). For the generalization problem, the second peak had the same amplitude as the initial peak. We attribute the first peak to decoding timed to the onset of the stimulus and the second peak to decoding timed to the cessation of the stimulus.

Decoding hue

The experiments enabled us to perform twelve hue identity tests (**Figure 2c left**) and twelve hue generalization-across-luminance-contrast tests (**Figure 2c right**). **Figure 2d** shows the average performance across the 12 identity problems (solid line) and the 12 generalization problems (dashed line; other conventions as in Figure 2b). The plot shows that classifiers could decode hue in both cases. Notably, the significance of the hue generalization-across-luminance-contrast tests supports the hypothesis that the brain has a representation of hue that is separate from the representation of luminance contrast. As with the luminance-contrast decoding problems, decoding performance had a higher peak for the identity problems (74% [71 77]) compared to the generalization problems (59% [56 63]), which provides additional support for the hypothesis that the brain also has a representation of hue that is inseparable from the representation of luminance contrast. The time to peak was not different for the identity and generalization problems (dashed vertical blue line, 125ms [115 130]; solid vertical blue line, 122ms [110 130]; p=0.193). The computed latency for the hue identity problem (75ms) was earlier than the hue generalization problem (92 ms; p=0.004). But the time at which the bootstrapped decoding problems rose above the upper 95% CI of baseline decoding was the same for the two problems (75 ms), suggesting that the generalization and identity problems have the same latency.

The graphs in **Figure 2** show that the time point of peak hue decoding corresponded to a dip in the luminance generalization decoding curve, and vice versa, the time point of peak luminance decoding corresponded to a notch in the hue generalization decoding curve (the vertical blue and gray lines in panels a and b are the same, for reference; see Figure 3 for quantification). In addition, the graphs show that the hue generalization problem was significant for a longer duration compared to the hue identity problem, which is a different outcome to the one observed for the luminance-contrast problems where the generalization problems were, on average, significant for a shorter duration than the identity problems. Finally, to the extent the graph in **Figure 2d** shows a second peak (open arrowhead) corresponding to the cessation of the stimulus, it was less pronounced than the initial decoding peak (curved arrow) (the open arrowhead is at the same y-axis level in both Figure 2c and Figure 2d to facilitate comparison).

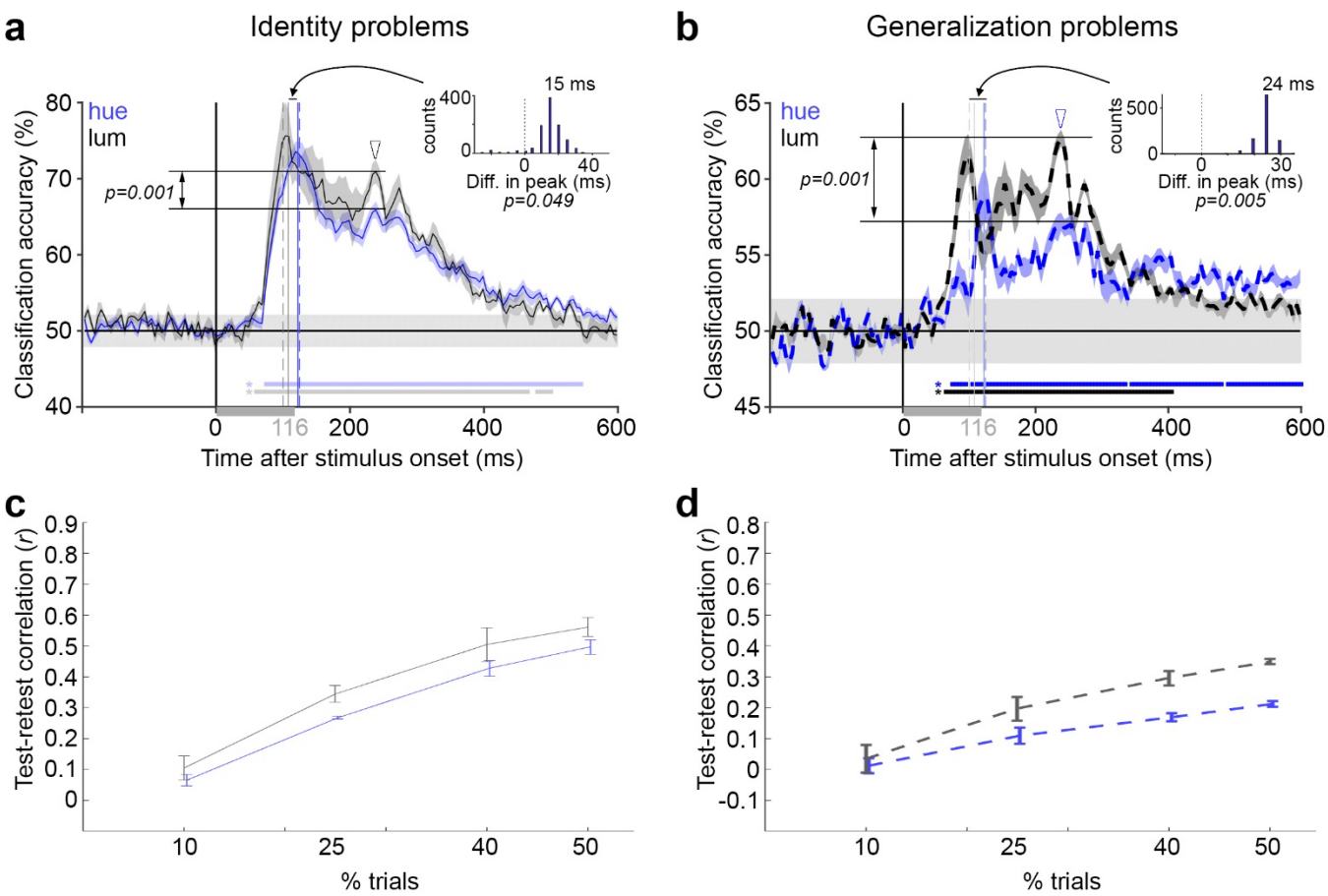

Figure 3. Comparing the temporal dynamics of decoding for luminance contrast and hue
a) Identity problems. Inset shows the differences between the peaks across the 1000 bootstrapped samples. Accuracy for luminance contrast peaked 15ms before hue. Classifier accuracy corresponding to the cessation of the stimulus showed a prominent peak for luminance contrast but not hue (open arrowhead). **b)** Generalization problems. Other conventions as Figure 2 b,d. **c)** Power analysis for identity problems. For each identity problem, pairs of classifiers were trained and tested on independent samples containing 10%, 25%, 40%, and 50% of the data, and the correlation

between the classifiers' performance at each time point was calculated (the analysis shows the extent to which the shape of the classification curve is similar for independent data sets of different sizes). This procedure was repeated five times to obtain the test-retest correlation and error bars. **d)** Power analysis for generalization problems. The plot was produced using the same procedure from **c**.

Figure 3 shows the two identity decoding curves (luminance and hue) on the same axes (**Figure 3a**), and the two generalization decoding curves on the same axes (**Figure 3b**). These plots underscore four main findings. First, representations of luminance contrast and hue did not follow the same time course. After the onset of significant decoding, as classification performance increased for the luminance generalization problem, classification performance decreased for the hue generalization problem, and vice versa (the black and blue curves are in counterphase in **Figure 3b**). This result can be quantified as the correlation of the derivative of the decoding curves, computed for 116 ms following the onset of significant decoding (116ms is the stimulus duration; R= -0.41 [-1.57 -0.11]). Second, hue was decodable after luminance contrast, as assessed by the time of peak decoding using either the generalization problems (24 ms delay, p=0.005) or identity problems (15 ms delay, p=0.049), and by the latency of decoding onset using either the generalization problems (p<0.001) or the identity problems (p=0.03). **Figure 3c,d** shows a power analysis to evaluate data reliability. The test-retest curves can only be computed using at most 50% of the data, so they underestimate the total power in the experiment. Third, for both the identity problems and the generalization problems, the second peak was larger for luminance contrast than for hue (double-headed arrow: p=0.001 for the identity problem; p=0.001 for the generalization problem). By comparison, the initial peak for hue was not different than the initial peak for luminance contrast (p=0.302 for the identity problem; p=0.121 for the generalization problem). The difference in peak between the hue and luminance decoding curves was different for the first peak versus the second peak, for both the identity problems ($p=10^{-68}$) and the generalization problems ($p=10^{-216}$). Fourth, hue was decodable for a longer duration compared to luminance, as assessed with either the identity problems or the generalization problems.

Greater cross-temporal decoding for hue than luminance contrast

The results discussed so far evaluate classifier performance using test data at the same time point after stimulus onset as the data used in training. The classifiers show significant decoding for a substantial amount of time. One possibility is that the pattern of activity is relativity stable over this time period; another possibility is that it is dynamic. To distinguish between these alternatives, we trained classifiers on the patterns of activity at each point in time and evaluated generalization to all other points in time. If

activity patterns are dynamic, the analysis will recover strong decoding performance only for situations in which the training and testing data sets were obtained at the same timepoint relative to stimulus onset, i.e. along the diagonal in a cross-temporal decoding plot. Alternatively, if activity patterns are relatively stable, the analysis will show strong decoding performance at time points away from the diagonal.

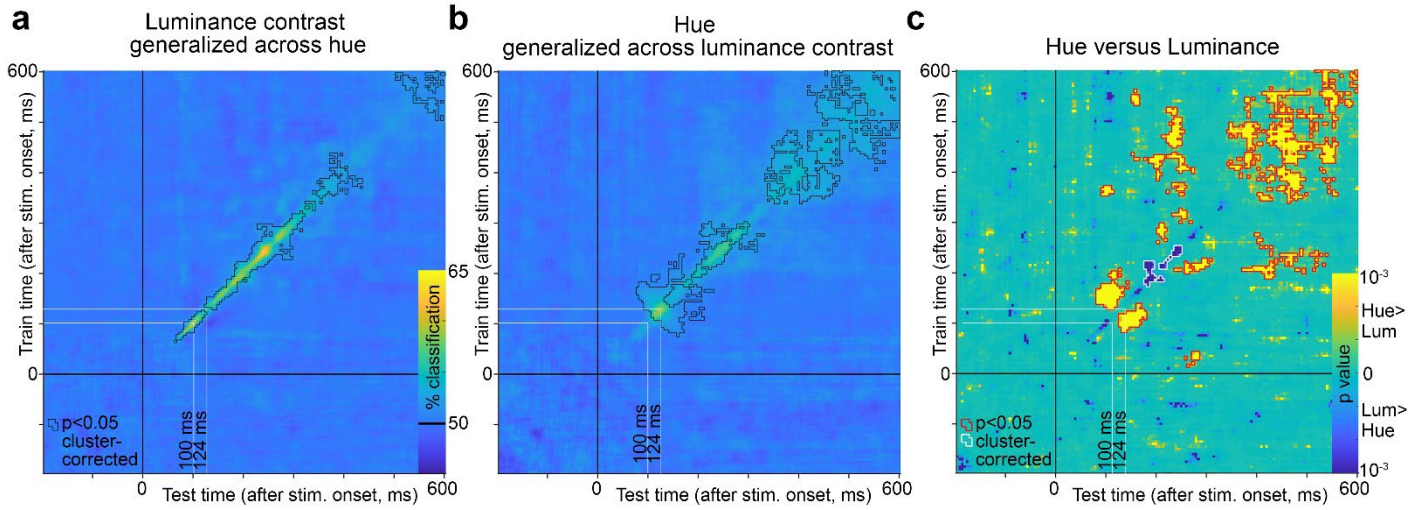

Figure 4. Testing the extent of cross-temporal generalization for decoding luminance contrast and hue
a) Classifiers were trained using the pattern of MEG activity elicited at time points from -200 ms to 600 ms after stimulus onset (y-axis) and tested using data not used in the training, across the same time interval, on the set of problems generalizing luminance contrast across hue. The best decoding performance was achieved for classifiers that were trained and tested using data from the same time point after stimulus onset, indicated by the strong performance along the x=y diagonal. The peak classification time was at 100 ms; there was a dip in classification performance at 124 ms. **b**) Data as in panel a, but for classifiers trained and tested on the set of hue generalization problems (hue invariant to luminance contrast). The peak classification time was 124 ms. The black contours show regions in the heatmap that were p<0.05 cluster corrected: they extended over a region of the map greater than any region of p<0.05 in the time bins before stimulus onset. **c**) Comparison of the classification performance for the luminance-contrast generalization problems (a) and the hue generalization problems (b). The time points where the classifiers were more accurate for luminance contrast compared to hue is shown as dark blue (white contours show cluster-corrected results), while the time points where the classifiers were more accurate for hue compared to luminance contrast is shown as yellow (red contours show cluster-corrected results). The p values were obtained by bootstrapping over problems and FDR corrected.

Figure 4 shows the cross-temporal decoding plots for the average of the generalization problems for decoding luminance contrast (**Figure 4a, left**) and hue (**Figure 4a, right**), averaged over the 12 individual problems for each. The color scale in the heatmap shows the percent classification—the values along the diagonal are the same as those shown in **Figure 3b**. The peak decoding performance for the luminance-contrast generalization problem was at 100 ms, while the peak decoding performance for the hue

generalization problem was at 124 ms (note the relatively lower decoding performance at 110 ms for the hue problem and at 124 ms for the luminance problem, corresponding to the counter-phase nature of the dashed curves in **Figure 3b**). The black contours identify data that were FDR and cluster-corrected to mitigate false positives attributed to multiple comparisons (clusters were defined as contiguous $p<0.05$ locations in the plot that were larger than any contiguous $p<0.05$ region from train and test time periods before stimulus onset). Decoding hue showed more cross-temporal generalization compared to decoding luminance contrast: successful decoding was evident further away from the diagonal in **Figure 4a** compared to **Figure 4b**. **Figure 4c** quantifies this comparison: yellow regions in the plot indicate where hue decoding was greater than luminance-contrast decoding, while dark blue regions indicate where luminance-contrast decoding was greater than hue decoding (the comparison was determined through a bootstrap procedure; red and white contours show cluster-corrected results). The relatively stronger cross-temporal generalization for hue, indicated in the plot as the yellow flanks around the diagonal, began at about the same time as the initial peak in hue generalization decoding (~124ms).

These results support the hypothesis that the patterns of activity in the brain associated with hue not only persist longer but also are more stable than the patterns of activity associated with luminance contrast, even though the overall peak decoding performance for luminance contrast (especially for the second peak corresponding to stimulus cessation) was higher than the peak decoding performance for hue. The relatively stronger cross-temporal generalization for hue compared to luminance contrast cannot be attributed to differences in the peak decoding performance because peak decoding for luminance contrast was, if anything, higher than for hue.

Luminance-contrast decoding varies with hue

The decoding analysis in **Figure 2** shows that luminance contrast can be decoded from the pattern of MEG data but does not address the variability in the extent to which luminance-contrast information can be decoded for different hues. We were interested in addressing this question because some behavioral data suggest that luminance is less reliably extracted from colors associated with the daylight locus (orange/blue) compared to the anti-daylight locus (pink/green), which has important implications for understanding the neural mechanisms that support color constancy (see Introduction). **Figure 5a,b** shows the individual luminance-contrast decoding problems averaged over participants; the plot reveals

substantial variability both in the four identity problems (thin gray lines) and the twelve generalization problems (thin blue lines). **Figure 5c** shows the peak decoding performance at any time following stimulus onset (0 to 600 ms) for the sixteen different problems. Data along the inverse diagonal correspond to the identity problems—problems in which the target hue carrying the luminance-contrast signal, in training the classifiers, was the same as the hue in the test data. Data off the diagonal correspond to the various generalization problems. The time bin at which peak decoding for each problem was achieved is indicated in each entry, along with the 95% CI of the peak decoding performance. The bolded numbers in each entry show the total number of 5ms time bins in which decoding was significant (corrected for false discovery rate, FDR). Entries in which significant decoding was not achieved for more than 5 consecutive time bins are indicated with an X. **Figure 5d** shows the decoding performance at the same time point for all problems: the time bin corresponding to the average peak decoding for the identity problem (105ms-110ms). The 95% CI of the classification performance is shown for each problem. **SI Movie 1** shows the decoding performance for all time bins -200 ms to 600ms after stimulus onset.

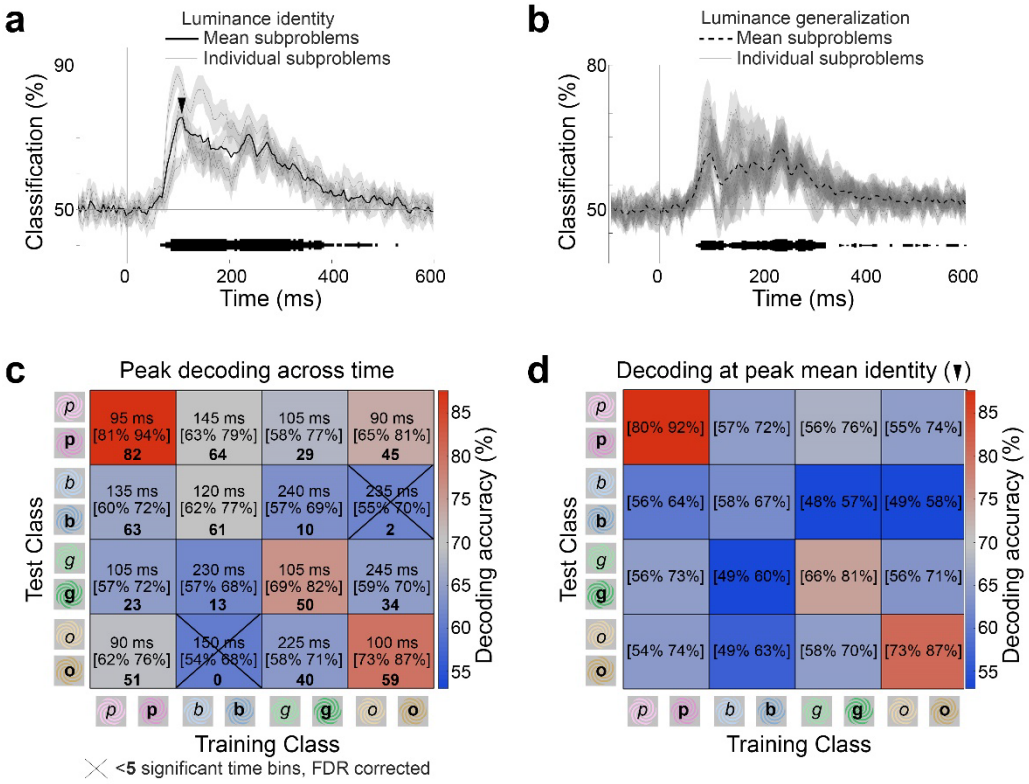

Figure 5. Decoding luminance contrast

a) Accuracy of classifiers trained to decode luminance contrast for stimuli of a given hue (4 problems, thin black lines; average, thick black line), and stimuli of different hues (12 problems, thin blue lines; average, thick blue line; shading shows standard error across the 18 subjects). The thickness of horizontal lines at the bottom show the number of decoding problems (of 4, left; of 12, right) that were significant in each 5ms bin. **b)** Heatmap showing the peak decoding accuracy for each of the 16 binary classifiers at the time point of peak decoding for each problem**.** Square brackets show the 95%CI of the classification accuracy. The time bin in which the classifier performance reached peak for each

problem is indicated in each entry in the heatmap. The bold numbers show the largest number of consecutive time bins with significant classification performance for each problem. **c)** Heatmap showing each classifier's performance at the time of its peak classification accuracy. The time to peak is indicated for each subproblem, and the 95% CI of the classification accuracy is shown in square brackets. The number of time bins where significant decoding was achieved is indicated in bold. Subproblems that were not significant for more than five consecutive bins are marked with an X. **d)** Heatmap showing the classifier performance at the same time point for all classifiers. The time point corresponds to the peak classification accuracy across the luminance-identity problems (105-110 ms after stimulus onset). The square brackets show the 95%CI of the classifiers. See **SI Movie 1** for the corresponding panels at all time bins -200 to +600ms after stimulus onset.

The data in **Figures 5c,d** support three findings. First, luminance-contrast decoding was strongest when the carrier (training) hue was the same as the test hue (unpaired t test on average peak decoding values comparing on-diagonal to off-diagonal **Figure 5c**, p=0.0005; **Figure 5d**, p=0.0014). Second, among these identity problems, classification performance was higher for warm colors (pink, orange) than cool colors (blue, green; paired t test across individual average warm color performance and individual average cool color performance, **Figure 5c**, p=0.007; **Figure 5d**, p=4e-5). Third, classifiers failed to achieve significant classification accuracy when training and testing across hues associated with the daylight axis: classifiers trained to distinguish luminance contrast of blue stimuli failed to decode luminance contrast of orange stimuli; and classifiers trained to distinguish luminance contrast of orange stimuli failed to decode luminance contrast of blue stimuli. Classifiers trained and tested using hues associated with the daylight axis (blue and orange) were less successful at decoding luminance than classifiers trained and tested using hues associated with the anti-daylight axis (pink and green; repeated measures one-way ANOVA, **Figure 5b**, p=0.03; **Figure 5c**, p=0.02).

fMRI-guided MEG source localization

We were interested in evaluating the extent to which MEG signals arising from functionally defined regions in the cortex could support decoding of hue and luminance contrast. To do so, we ran fMRI experiments in 14 of the same participants in whom we collected MEG data and performed the MEG analyses using subject-specific source localization. Our goal was to use functional data to define regions of interest in the ventral visual pathway in each participant, controlling for individual differences in the absolute location of functional domains across people. In each subject we used fMRI to identify regions biased for faces, places, colors, and objects, using the same paradigm we used previously in which we measured fMRI responses to short movie clips of faces, bodies, objects, and scenes (Lafer-Sousa et al. , 2016). The paradigm involved measuring responses to intact and scrambled versions of the clips, and to clips in full color and black-and-

white. As described in Lafer-Sousa et al (2016), the results allow one to define a set of regions of interest of the ventral visual pathway: face-biased regions (including the FFA); place-biased regions (including the PPA); and color-biased regions sandwiched between the face-biased and place-biased regions. The results also recover area LO, defined by stronger responses to intact versus scrambled movie clips.

Figure 6a shows the fMRI results for one participant: greater responses to colored movie clips compared to black-and-white versions of the movie clips is shown by the heat map; functional domains for faces (faces>objects, including the FFA), objects (intact objects>scrambled objects, LO), and places (places>objects, including the PPA) are indicated by contour maps drawn at p=0.001 threshold. Color-biased activity was found sandwiched between place-biased activity (medially) and face-biased activity (laterally), confirming prior observations (Lafer-Sousa et al., 2016). By aligning each participant to a standard atlas (Desikan et al. , 2006, Toga et al. , 2006) we also generated regions of interest for V1, V2, and MT, and for frontal cortex and the precentral gyrus (control regions).

MEG signals source localized to V1 and V2 yielded the highest magnitude current source density averaged across all stimulus presentations (**Figure 6b**, left panel). The magnitude of the CSD was different among the functional regions identified in the ventral visual pathway (p=0.002, repeated measures one-way ANOVA): the color-biased regions were not different from the FFA (p=0.12; paired t-test); but were different from LO (p=0.01), and from PPA (p=0.02) (**Figure 6c**). These results provide a direct measure of neural activity, and confirm the indirect measurements obtained with fMRI suggesting that fMRI-identified color-biased regions (and possibly face-biased regions) play an important role in color processing.

Luminance contrast generalized across hue was decodable to some extent in all visual regions except the face-biased regions and the color-biased regions; it was most decodable in V1 and V2, and to a lesser extent in MT and LO; and to an even lesser extent in the place-biased regions; it was not decodable in the two control regions (**Figure 6d**). Hue generalized across luminance contrast was not decodable in any region except to a very small extent in V2 (**Figure 6e**). The distribution of sensors used in the decoding analysis is shown in **Figure 6f,g** (see legend and methods for details).

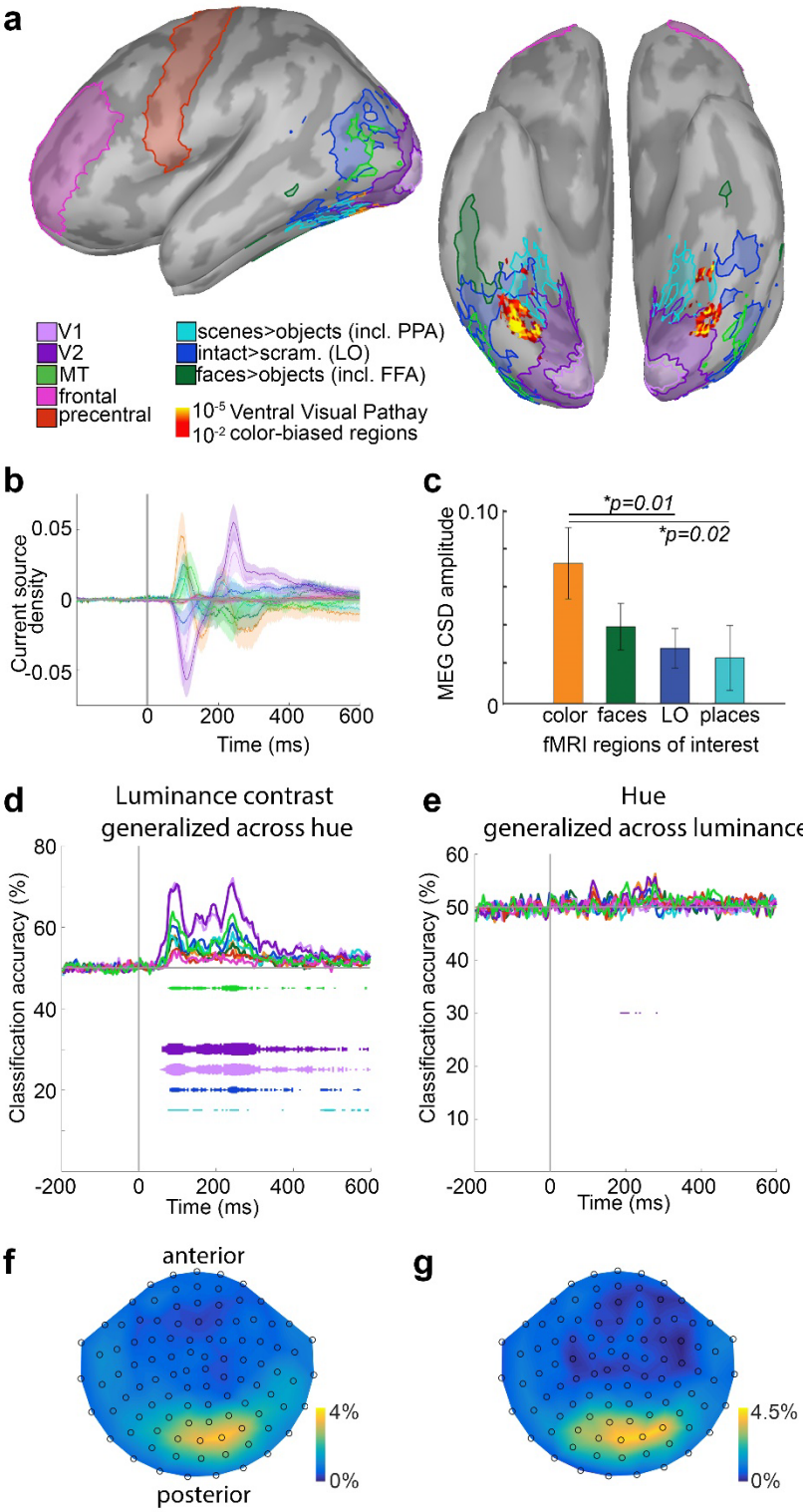

Figure 6. Source localization to regions defined using fMRI

a) Example anatomy from one participant with regions of interest mapped on the inflated cortex. Functional regions were defined using movie clips in color and black-and-white of faces, bodies, objects, scrambled objects using the same procedure as in Lafer-Sousa et al (2016). The activation map shows voxels with higher responses to color clips compared to black-and-white clips. Contours show regions of interest for faces > objects (including FFA), intact shapes > scrambled shapes (LO), and places > objects (including PPA). Regions of interest for V1, V2, MT, frontal, and precentral ROIs were defined anatomically. **b)** Relative current source density (CSD) of each ROI over time in response to the spiral stimuli, averaged across participants and calculated with dynamical Statistical Parametric Mapping (dSPM; see Methods). 0 on the x-axis is stimulus onset. Values on the y-axis are unitless. Transparent shading shows SEM. **c)** Amplitude of CSD in each ROI, calculated as distance from peak to trough of the time course in panel b. Error bars are SEM. There was a significant effect of ROI on response magnitude (repeated measures one-way ANOVA, p = 0.002). Responses source-localized to the color-biased regions were significantly different from those source-localized to LO (paired t-test, p=0.01) and place-biased regions (p=0.02), but not face-biased ROIs (p=0.12). **d)** Average classifier performance on the luminance contrast generalized over hue problems (12 problems averaged together; see Figure 2a) trained using only those MEG data localized to the MRI-defined ROIs (N=14 participants). Each line shows the average accuracy of one ROI-restricted classifier averaged across participants (color key in panel a). **e)** Average classifier performance on the hue generalized ove luminance contrast problems (12 problems averaged together (see Figure 2b). Other conventions as for panel d. **f)** The distribution of sensors used as features for the classifiers across participants (N=18). Color bar shows the percent likelihood that any given sensor was selected as a feature. **g)** As in panel (f), but for decoding hue generalizing across luminance contrast.

Discussion

The experiments presented here used multivariate analyses of MEG responses to carefully calibrated color stimuli, and produced three new findings: first, hue and luminance contrast could be decoded independently, and with different timing. Classification accuracy for luminance contrast peaked 100-108 ms after stimulus onset, while accuracy for hue peaked a little later, 122-125 ms after stimulus onset. Second, representations of hue showed greater cross-temporal generalization and were more sustained than representations of luminance contrast. And third, representations of hue and luminance contrast showed some interaction, and these interactions varied systematically with hue: luminance signals attached to orange and blue (colors associated with the daylight locus) were decoded less reliably than luminance signals associated with pink and green (colors associated with the anti-daylight locus). These results have implications for our understanding of the different roles that luminance contrast and hue play in visual perception.

Multivariate analyses of signals acquired across the brain, such as with MEG, provide a powerful tool to uncover the way in which perceptual experiences are encoded (Haynes and Rees, 2006, Tong and Pratte, 2012); and multivariate analyses specifically of MEG data uncover important information about the time taken by the brain to perform computations (Carlson et al., 2013, van de Nieuwenhuijzen et al., 2013). In the present work, the difference in decoding time for luminance contrast compared to hue suggests either these attributes are encoded to some extent by different neural populations, or if these attributes are encoded by entirely the same population, the encoding must involve temporal multiplexing. Either possibility argues against the notion that luminance contrast and hue are always multiplexed completely simultaneously by the same neural population.

In object perception, decoding latency reflects the perceptual and categorical dissimilarity of stimuli, with more perceptually dissimilar stimuli and abstract categories decodable later, and associated with computations performed by areas further along the visual-processing hierarchy (Proklova et al., 2019, Carlson et al., 2013, Cichy et al., 2014) (an example of increasing category abstraction is Doberman→dog→animal→animate). One way of thinking about the relatively later decoding of hue, then, is that (1) hue discrimination involves greater perceptual dissimilarity or greater category abstraction than does luminance-contrast, and (2) it is computed either by circuits downstream of those that compute luminance contrast or requires more recurrent processing than computations of luminance contrast. The

time course for decoding hue (peak 124ms; **Figure 3**) is comparable to that for decoding shape-independent object category (Kaiser et al. , 2016) and face identity (Dobs et al. , 2019), operations that probably reflect activity in area LO and the fusiform face area (FFA). The decoding time course for hue therefore raises the possibility that hue is computed at about the same distance along the visual processing hierarchy as LO and the FFA, which implicates the posterior color-biased region of the ventral visual pathway (Lafer-Sousa and Conway, 2013, Lafer-Sousa et al., 2016)—in some accounts, this region is included as part of the V4 Complex (Bannert and Bartels, 2018). Within this region there are compartments comprising neurons that are spatially organized according to hue, and whose hue selectivity is tolerant to changes in luminance contrast (Bohon et al. , 2016, Conway, 2009), consistent with the notion that these neurons represent hue in a way that is tolerant to changes in luminance contrast.

Neural representations related to object vision that emerge earliest, as determined by classifiers trained on specific exemplars, reflect the encoding of "low-level visual features"(Carlson et al., 2013)—decoding in these classifiers peaks fairly early, at ~100ms, and is attributed to operations implemented early in the visual-processing hierarchy, perhaps V1 (Cauchoix et al. , 2014, Goddard et al. , 2016, Kaiser et al., 2016, Martin Cichy et al. , 2017). But it is not clear what these visual features comprise. It is often implied that low-level features include oriented luminance-contrast edges and color, for example because these are the features extracted in early layers of convolutional neural networks (Krizhevsky et al., 2012). But the relatively earlier timing of luminance-contrast decoding compared to hue decoding, which peaked at about 100 ms, suggests that (1) luminance contrast (and not hue) may be among the first features to be encoded; and (2) luminance contrast is implemented at an early stage in the visual-processing hierarchy, before hue, perhaps in V1.

What role do luminance contrast and hue play in vision? Under normal conditions, the visual system is confronted with a constant stream of retinal images, each of which is associated with a cascade of neural activity lasting a second or longer (Marti and Dehaene, 2017). On the one hand, presumably the visual system must rapidly parse this stream of information to enable encoding of new content. On the other hand, high-level components of the visual system likely retain some representations for longer durations to enable recognition and memory. Analyzing the timing differences in decoding, including the extent to which representations generalize across time, may provide clues to how the visual system achieves these apparently competing objectives (King and Dehaene, 2014). Regarding the first objective: In any situation in

which information is encoded in time, it is advantageous to have clear signals indicating the start and end of the code, such as in genetics where gene sequences are parsed by canonical start and stop codons. Temporal sequences of brain activity may play a similar role, for example indicating the initiation and termination of action sequences in nigrostriatal circuits (Jin and Costa, 2010). In object vision, dynamical systems modeling predicts the existence of observable update mechanisms that signal new content (King and Wyart, 2019). In the present work, the time course over which luminance-contrast could be decoded had clear peaks corresponding to both the onset and cessation of the stimulus (**Figures 2, 3**). Moreover, representations of luminance contrast showed very limited cross-temporal generalization, forming a narrow band confined to the identity diagonal in **Figure 4a**. The time course for decoding hue, meanwhile, showed a clear peak only at stimulus onset, was more sustained, and showed relatively stronger cross-temporal generalization (**Figure 4b**). This pattern of results shows that the representation instantiated by luminance contrast is well defined in time, which may reflect the temporal precision of lateral geniculate neurons (Reinagel and Reid, 2002). The pattern of results is therefore consistent with the idea that the brain uses luminance-contrast signals and not hue, as the updating signal, to encode discrete events embedded in the constant stream of visual information.

The representation of hue not only had a less prominent peak signaling stimulus cessation, but was also more sustained and showed greater cross-temporal generalization compared to the representation of luminance contrast, which implies that the neural representation instantiated by hue is more stable than the one instantiated by luminance contrast. The greater cross-temporal generalization for hue began much earlier (~124 ms) than the cross-temporal generalization attributed to differences in task performance or imagery as observed by others, which probably reflect top-down feedback (Andersen et al. , 2016, Dijkstra et al. , 2018, Hebart et al. , 2018, Marti and Dehaene, 2017, Marti et al. , 2015, Quentin et al. , 2019). Classifiers trained on real colors successfully decode stimuli with implied color only 200 ms after stimulus onset, providing a benchmark for the timing associated with high-level cognition on decoding color (Teichmann et al. , 2019). The relatively earlier cross-temporal generalization for hue reported here suggests that it derives from a feed-forward representation: something about the way color is initially encoded by the visual system is associated with a more stable representation compared to luminance contrast. This pattern of results provides a neural correlate for the preeminence of color over luminance contrast in object grouping, as evident in the difficulty one faces in making heterochromatic brightness matches. The relative stability of hue, compared to luminance contrast, is evident in the difficulty people

with normal color vision face in making heterochromatic brightness matches, a difficulty that is exploited as a control plate in the famous Ishihara test of color vision defects (numbers demarcated by luminance contrast that are imbedded in colored noise are hard to see). Taken together, the timing differences in the representations for hue versus luminance contrast provide a neural correlate for the relatively different roles that these stimulus dimensions play in perception: luminance contrast tends to be used by the visual system to constantly update representations of scene structure; while representations of hue linger and are useful in grouping and remembering visual information (Gegenfurtner and Rieger, 2000).

While the results show that representations of hue and luminance contrast are somewhat decoupled, they also indicate that representations of hue and luminance contrast must be yoked to some extent: decoding performance for the identity problems, in which classifiers were trained using responses to stimuli that were distinguished by a combination of luminance contrast and hue, was always better than decoding performance for the generalization problems, in which classifiers were trained using responses to stimuli that were only distinguishable by one dimension, invariant to the other (**Figure 2**). These results are consistent with the idea that encoding of these dimensions is multiplexed simultaneously by some population of neurons in both the geniculate (Wiesel and Hubel, 1966) and V1 (Gegenfurtner, 2003, Horwitz and Hass, 2012, Johnson et al. , 2008). The timing of decoding is consistent with the hypothesis that the yoked representation derives from activity early in the visual-processing hierarchy (Carlson et al., 2013), perhaps in subcortical circuits.

Luminance information was not carried equally by all hues, which provides clues to the neural mechanisms that support color constancy: classifiers trained to distinguish light and dark blue were incapable of distinguishing light and dark orange; and classifiers trained to distinguish light and dark orange were incapable of distinguishing light and dark blue (**Figure 5**). These results reflect an asymmetry in the representation of colors of the daylight axis (orange-blue) versus colors of the anti-daylight axis (pink-green). The results show that luminance is not represented entirely independently of hue, and they support the idea that the neural representation is adapted to natural lighting conditions. Under natural viewing conditions, the chromaticity of the illuminant, which is restricted to oranges (direct sunlight) and blues (indirect skylight), covary with luminance contrast (shadows, which are dark, reflect the sky, which is relatively blue). The results predict that luminance information arising from an orange or blue surface is not reliable, which has implications for models of color constancy. The results may also explain why the specific

color combination of the famous #thedress image brings about the essential ambiguity of the image (Lafer-Sousa and Conway, 2017, Lafer-Sousa et al., 2015, Winkler et al., 2015).

One chief advantage of MEG over other common non-invasive techniques for measuring brain function such as fMRI is that MEG signals are directly attributable to neural events. By contrast, fMRI measures blood flow which is indirectly related to neural activity. There are substantial gaps in knowledge regarding the connection between fMRI and neural events. One disadvantage of MEG is the relatively low spatial resolution. To leverage the high spatial resolution of fMRI on the one hand and the more direct access to neural events of MEG on the other hand, we used both techniques in the same participants, exploiting source localization to estimate MEG signals arising from fMRI-identified regions defined in each individual subject (**Figure 7**). The results provide a way of independently testing conclusions drawn from fMRI experiments. Within the ventral visual pathway, cortical regions showing the strongest fMRI responses to color are sandwiched between more lateral regions responding most strongly to faces and more medial regions responding most strongly to places (Lafer-Sousa et al., 2016); this pattern is also seen in macaque monkeys (Lafer-Sousa and Conway, 2013). Among regions of the VVP, the MEG signals assigned to color-biased regions showed the largest current-source density in response to the stimuli used in the MEG experiments—these stimuli differed only in color and not shape. Thus the present results support the idea that the VVP comprises parallel streams characterized by differential sensitivity to color information (Conway, 2018). Source-localized analyses did not recover significant luminance-invariant hue representations, which may not be surprising given fundamental limitations of source localization (Cicmil et al. , 2014).

Methods

Visual Stimuli

Stimuli were eight square-wave spiral gratings on a neutral gray background (**Figure 1a**) (Brouwer and Heeger, 2009, Mannion et al., 2009, Seymour et al., 2010). The 8 stimulus colors, four hues at two luminance-contrast levels, of matched cone contrast, were defined in DKL color space (Derrington et al., 1984, MacLeod and Boynton, 1979) using implementations by Westland (Westland et al. , 2012) and Brainard (Brainard, 1996): the axes of this color space are defined in terms of activation of the two cone-opponent post-receptoral chromatic mechanisms (**Figure 1a**). The z-axis is defined by luminance. The four hues were defined by the intermediate axes of DKL space: at 45° (pink), 135° (blue), 225° (green), and 315° (orange). Two spirals – one high luminance (20° elevation; "light") and one low luminance (340° elevation; "dark") -- were created at each hue. The neutral adapting background was 50 cd/m2. The luminance contrast of the stimuli was 26%. Modulation of the cone-opponent mechanisms, shown in **Figure 1c,d**, was computed relative to the adapting background gray, using the Stockman and Sharpe 2degree cone fundamentals, Judd corrected.

MEG Acquisition and Preprocessing

Participants were scanned in the Athinoula A. Martinos Imaging Center of the McGovern Institute for Brain Research at the Massachusetts Institute of Technology (MIT) over the course of 2 sessions, on an Elekta Triux system (306-channel probe unit consisting of 102 sensor triplets, with 204 planar gradiometer sensors, and 102 magnetometer sensors). Stimuli were back-projected onto a 44" screen using a SXGA+ 10000 Panasonic DLP Projector, Model No. PT-D10000U (50/60Hz, 120V). Data was recorded at a sampling rate of 1000Hz, filtered between 0.03-330Hz. Head location was recorded by means of 5 head position indicator (HPI) coils placed across the forehead and behind the ears. Before the MEG experiment began, 3 anatomical landmarks (bilateral preauricular points and the nasion) were registered with respect to the HPI coils, using a 3D digitizer (Fastrak, Polhemus, Colchester, Vermont, USA). During recording, pupil diameter and eye position data were collected simultaneously using an Eyelink 1000 Plus eye tracker (SR Research, Ontario, Canada) with fiber optic camera.

Once collected, raw data was preprocessed to offset head movements and reduce noise by means of spatiotemporal filters (Taulu et al, 2004; Taulu & Simola, 2006), with Maxfilter software (Elekta,

Stockholm). Default parameters were used: harmonic expansion origin in head frame = [0 0 40] mm; expansion limit for internal multipole base = 8; expansion limit for external multipole base = 3; bad channels omitted from harmonic expansions = 7 s.d. above average; temporal correlation limit = 0.98; buffer length = 10 s). In this process, a spatial filter was applied to separate the signal data from noise sources occurring outside the helmet, then a temporal filter was applied to exclude any signal data highly correlated with noise data over time. Following this, Brainstorm software (Tadel et al., 2011) was used to extract the peri-stimulus MEG data for each trial (-200 to 600 ms around stimulus onset) and to remove the baseline mean.

MEG Participants and Task

All participants (N=18, 11 female, age 19-37 years) had normal or corrected-to-normal vision, were right handed, and spoke English as a first language. One participant was an author and thus not naïve to the purpose of the study. During participants' first session, they were screened for colorblindness using Ishihara plates; they also completed a version of a color-naming task as part of a separate study. After this task, participants completed a 100-trial practice session of the 1-back task that would be used in the MEG experimental sessions. Once this was complete, participants were asked if they had any questions about the task or the experiment; eye-tracking calibration was performed; and MEG data collection began.

In the 1-back task, participants were instructed to fixate at the center of the screen. Spirals were presented subtending 10° of visual angle, for 116 ms, centered on the fixation point, which was a white circle that appeared during inter-trial intervals (ITIs, 1s). In addition to the spirals, the words "green" and "blue" were presented in white on the screen for the same duration, and probe trials were presented with a white "?". Responses to the words were analyzed as part of a separate study. During the probe trials, which occurred every 3-5 stimulus trials (pseudorandomly interspersed, 24 per run), participants were instructed to report via button press if the two preceding spirals did or did not match according to hue (1-back hue task). Maximum response time was 1.8s, but the trials advanced as soon as participants answered.

Participants were encouraged to blink only during probe trials, as blinking generates large electrical artifacts picked up by the MEG. Each run comprised 100 stimulus presentations, and participants completed 25 runs per session over the course of approximately 1.5 hours. Between each run, participants

were given a break to rest their eyes and speak with the researcher if necessary. Once 10s had elapsed, participants chose freely when to end their break by button-press. Over the course of both sessions, participants viewed each stimulus 500 times.

In addition to the 18 participants analyzed in the main thrust of this study, a pilot version was deployed with 2 participants (1 female, age 20-30 years) to determine the behavioral task and decoding parameters. This study differed from the main experiment in that participants completed 5 sessions of 20 runs each, and during each session, one half of the runs required the participant to perform the 1-back hue task, and the other half of the runs required the participant to match the two previous spirals according to luminance. The data from these participants was used to choose the parameters for the decoding analysis used in the rest of the study (see below). Additionally, the decoding results from these two participants showed no difference between MEG data collected in the hue-matching condition and the luminance-matching condition (**Figure S1)**, so only the hue-matching condition was used for the main experiment. Data from all participants was used (no data was excluded because of poor behavioral performance) .

All experimental procedures involving participants tested in laboratory were approved by the Wellesley College Institutional Review Boards, the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects, and the National Institutes of Health Intramural Institute Clinical Research Review Committee.

MEG Processing and Decoding Analyses

Brainstorm software was used to process MEG data. Trials were discarded if they contained eyeblink artifacts, or contained out-of-range activity in any of the sensors (0.1-8000 fT). Three participants exhibited sensor activity consistently out of range, so this metric was not applied to their data as it was not a good marker of abnormal trials. After excluding bad trials, there were at least 375 good trials for every stimulus type for every participant. Data were subsampled as needed to ensure the same number of trials per condition were used in the analysis.

Decoding was performed using the Neural Decoding Toolbox (NDT) (Meyers, 2013). We used the maximum correlation coefficient classifier in the NDT to train classifiers to associate patterns of MEG activity across

the sensors with the visual stimuli presented. This classifier computes the mean population vector for sets of trials belonging to each class in the training data and calculates the Pearson's correlation coefficient between those vectors and the test vectors. The class with the highest correlation is the classifier's prediction. The main conclusions were replicated when using linear support vector machine classifiers. The classifiers were tested using held-out data—i.e. data that was not used in training. Data from both magnetometers and gradiometers were used in the analysis, and data for each sensor was averaged into 5-ms non-overlapping bins from 200 ms before stimulus onset to 600 ms after stimulus onset.

Custom MATLAB code was used to format MEG data preprocessed in Brainstorm for use in the NDT and to combine the two data-collection sessions for each participant. Decoding was performed independently for each participant, and at each time point. As illustrated in **Figure 1**, for each decoding problem, at each timepoint (a 5 ms time bin), the 375 trials for each stimulus condition were divided into 5 sets of 75 trials. Within each set, the 75 trials were averaged together. This process generated 5 cross-validation splits: the classifier was trained on four of these sets, and tested on one of them, and the procedure was repeated five times so that each set was the test set once. This entire procedure was repeated 50 times, and decoding accuracies reported are the average accuracies across these 50 decoding "runs". This procedure ensured that the same data was never used for both training and testing, and it also ensured the same number of trials was used for every decoding problem. The details of the cross-validation procedure, such as the number of cross-validation splits, were determined during the pilot experiments to be those that yielded a high signal-to-noise ratio (SNR) and high decoding accuracy in both participants on the stimulus identity problem.

On each run, both the training and test data were z-scored using the mean and standard deviation over all time of the training data. Following others, we adopted a de-noising method that involved selecting for analysis data from the most informative sensors (Isik et al., 2014); we chose the 25 sensors in the training data whose activity co-varied most significantly with the training labels. These sensors were identified as those with the lowest p-values from an F-test generated through an analysis of variance (ANOVA); the same sensors were then used for both training and testing. The sensor selection was specific for each participant. The sensors chosen tended to be at the back of the head (**Figure 7f,g**). Analyses using all channels, rather than selecting only 25, yielded similar results.

All classification problems were binary (see **Figure 2**). For each problem illustrated in Figure 2, a classifier was trained and tested in 5ms bins from time t=200ms before stimulus onset to t=600ms after stimulus onset (see **Figure 1e**). The classifiers' performance shown in Figures 2 and 3 were generated through a bootstrapping procedure. First, the problems were evaluated for each participant (resulting in 18 independent decoding time courses). Then, for each unique problem, we averaged the decoding time courses across participants. The sets of identity and generalization problems, for hue and luminance contrast, were sampled, with replacement, 1000 times to generate the decoding traces in Figures 2 and 3. The gray shading shows the standard error of the bootstrap mean. We computed the 95% CI around the baseline, and significant decoding was defined as time bins in which decoding accuracy was higher than the upper bound of the 95% CI limit for five consecutive time bins. Onset of significance was calculated as the first time point where accuracy was significant for five continuous 5-ms time bins—the requirement that the accuracy be significant for five consecutive bins was adopted to minimize false positives, which ensured that spurious correlations in the baseline period were not marked as significant.

The power analysis in Figure 3 was obtained by drawing pairs of independent samples of 10%, 25%, 40%, and 50% of the trials (from a total of 375 trials), determining the correlation of the classification performance among the subproblems between the pairs, and repeating the procedure 5 times to generate error bars. For example, for the "10%" data point in the graph, we drew two sets of 10% of the trials at random—no trials were common to both sets. We trained separate classifiers on each of the independent sets. We correlated the classification performance at each 5ms time point for the results of these two sets of problems. We repeated this procedure 5x. We then averaged the correlation coefficients from the 5 repeats to obtain error bars.

In Figure 4, we tested the performance of the classifiers across time: each classifier trained using data obtained at each time bin was tested using data obtained at every 5 ms time bin from -200 to 600 ms after stimulus onset creating a 2-dimensional matrix of decoding results.

In Figure 5, we used a permutation to determine when classifier accuracies across participants were significantly different from chance (Pantazis et al, 2005). This was done by permuting the sign of the decoding accuracy data on a participant basis 1000 times. For each permutation sample, the mean accuracy was recomputed, resulting in an empirical distribution of 1000 mean accuracies. This distribution was used

to convert the real mean accuracies across subjects over time to p-value maps over time. The p-values assigned to each time point were then corrected to account for false discovery rate (FDR; Benjamini & Hochberg, 1995; Yekutieli & Benjamini, 1999). FDR-corrected p-values lower than 0.05 were taken as significant. The 95%CI of the classification performance for each problem (the entries in Figure 5b, 5c) were generated by bootstrapping across participants (n=1000).

Decoding analyses were also performed using eye tracking data collected during the MEG sessions. Two analyses were conducted: one using pupil diameter and one using eye position (**Figure S2)**. All parameters were identical to the MEG analysis except for the number of input features to the classifier. Rather than MEG sensors, the classifier used either the diameters of the two pupils (two features) or the xy coordinates of the positions of the two eyes (four features).

MRI dynamic Localizer Task

To localize shape, place, face, and color-biased regions of interest (ROIs), 14 of 18 participants were scanned using the fMRI dynamic Localizer (DyLoc) described in Lafer Sousa et al, 2016, with the same parameters described there. In brief, participants passively viewed full color and grayscale (achromatic) versions of natural video clips that depicted faces, bodies, scenes, objects, and scrambled objects. Scrambled objects clips were clips in the object category that were divided into a 15 by 15 grid covering the frame, the boxes of which were then scrambled. Participants completed 8 runs of the task, each of which contained 25 blocks of 18 s (20 stimuli and 5 gray fixation blocks). The stimuli were a maximum of 20° of visual angle wide and 15° tall. A Siemens 3T MAGNETOM Prisma fit scanner (Siemans AG, Healthcare, Erlangen, Germany) with 64 RF receivers in the head coil was used to collect MRI data in 8 of 14 participants, while a Siemens 3T MAGENTOM Tim Trio scanner with 32 channels in the head coil was used for the other 6 subjects.

For both groups, following Lafer-Sousa et al, a T2*-weighted echo planar imaging (EPI) pulse sequence was used to detect blood-oxygen-level-dependent (BOLD) contrast. Field maps (2 mm isotropic, 25 slices) were collected before each dyLoc run for the purpose of minimizing spatial distortions due to magnetic inhomogeneities in the functional volumes during analysis. Functional volumes (2 mm isotropic, 25 slices,

field of view [FOV] = 192 mm, matrix = 96x96mm, 2.0 s TR, 30 ms TE, 90˚ flip angle, 6/8 echo fraction) were collected on a localized section of the brain, aligned roughly parallel to the temporal lobe. The volumes covered V1-V4 in occipital cortex as well as the entirety of the temporal lobe ventral to the superior temporal sulcus (STS), and in some cases including parts of the STS. To allow for T1 equilibration, in each run, the first 5 volumes were not used during analysis.

High-resolution T1-weighted anatomical images were also collected for each subject by means of a multiecho MPRAGE pulse sequence (1 mm isotropic voxels, FOV = 256 mm, matrix = 256x256mm).

MRI Analysis

MRI data were processed following Lafer-Sousa et al (2016). Using Freesurfer (http://surfer/nmr/mgh.harvard.edu) and custom MATLAB scripts, the anatomical volumes were segmented into white- and gray-matter structures (Dale et al, 1999; Fischl et al, 1999, 2001). Functional data, processed on an individual subject basis, were field- and motion-corrected (by means of rigid-body transformations to the middle of each run), normalized for intensity after masking non-brain tissue, and spatially smoothed with an isotropic Gaussian kernel (3 mm FWHM) for better SNR. Subsequently, Freesurfer's bbregister was used to generate a rigid-body transformation used to align the functional data to the anatomical volume.

Whole-volume general linear model-based analyses were performed for all 8 runs collected for each participant, using boxcar functions convolved with a gamma hemodynamic response function as regressors (Friston et al, 1994); each condition's boxcar function included all blocks from that condition, as well as nuisance regressors for motion (three translations, three rotations) and a linear trend to capture slow drifts.

Brain regions used to restrict decoding analyses of MEG source data were defined using two methods. Anatomically defined regions were defined using surface-based Freesurfer atlases: "precentral" and "frontal" regions corresponded respectively to the "precentral" and "rostralmiddlefrontal" bilateral regions in the Desikan-Killiany atlas (Desikan et al, 2006); V1 (BA 17), V2 (BA 18) and MT regions were defined using the Brodmann atlas (http://ftp.nmr.mgh.harvard.edu/fswiki/BrodmannAreaMaps; Brodmann, 1909).

Functionally defined regions were defined individually using Lafer-Sousa et al (2016) as a reference. FFA was selected from voxels where face response>object response ($p<0.001$), using data from all 8 runs. The same procedure was followed for VVP-c from voxels where color response>grayscale response, PPA from voxels where scene response>object response, and LO from voxels where object response>scrambled object response.

Source Localization and Decoding with ROIs

Current source density is a metric representing the current at each point on the surface of the brain, defined by the source grid. First, using Brainstorm, a minimum norm estimate (MNE) was calculated, which was "depth-weighted", to compensate for a bias in current density calculations that results in more activity being placed on superficial gyrii, neglecting regions of cortex embedded in deeper sulci (Hämäläinen, 2009). The MNE at a given source was normalized by the square root of a local estimate of noise variance (dynamical Statistical Parametric Mapping; dSPM, Dale, 2000), yielding a unitless z-scored statistical map of activity. Once a source map was created, ROI analysis was performed by restricting the features of the classifiers to the top 25 sources within the bounds of a given ROI whose activity covaried most with the training labels, using custom code. Additionally, the sources within an ROI were averaged together within subjects to yield the average sensor response by ROI.

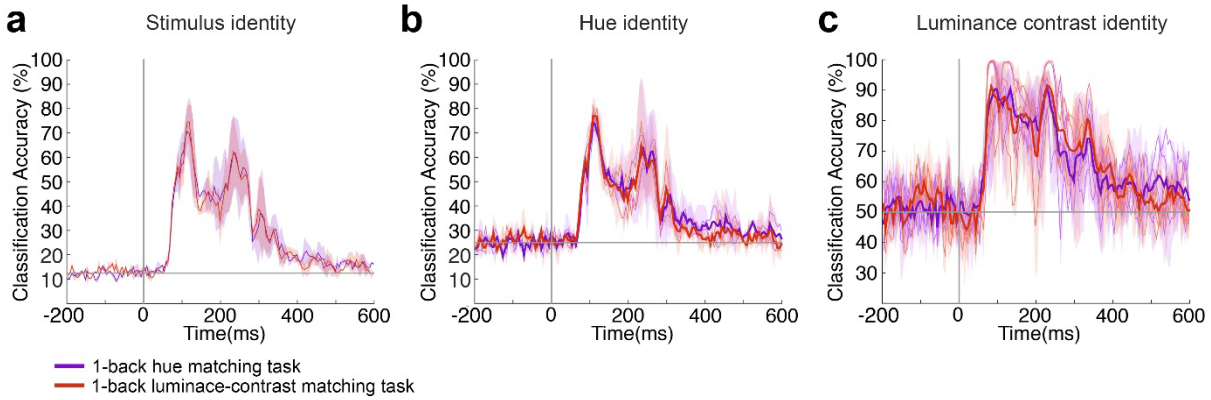

Figure SI1. Pilot-study results
a) Classifier performance decoding stimulus identity (chance is 1/8) in experimental sessions in which participants performed a 1-back hue-matching task (purple trace) or a 1-back luminance-matching task (red trace). The pilot experiment was done with 2 participants (1 female, age 20-30 years) who completed 5 sessions of 20 runs each, and during each session, one half of the runs required the participant to perform the 1-back hue task, and the other half of the runs required the participant to match the two previous spirals according to luminance. **b)** Classifier performance decoding hue identity (chance is 1/4). **c)** Classifier performance decoding luminance identity (chance is 1/2).

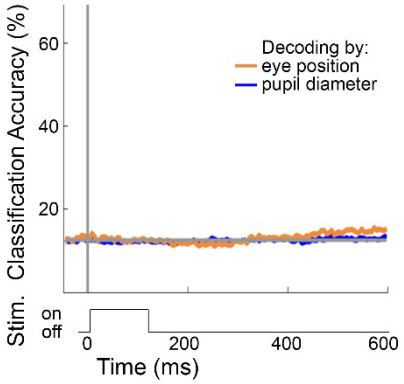

Figure SI2. Performance of classifier's trained on eye position and pupil diameter.
Classifiers trained to decode stimulus identity on eye position and pupil diameter (chance is 1/8).

References

Andersen LM, Pedersen MN, Sandberg K, Overgaard M. Occipital MEG Activity in the Early Time Range (<300 ms) Predicts Graded Changes in Perceptual Consciousness. Cereb Cortex. 2016;26:2677-88.

Bannert MM, Bartels A. Human V4 Activity Patterns Predict Behavioral Performance in Imagery of Object Color. Journal of Neuroscience. 2018;38:3657-68.

Bohon KS, Hermann KL, Hansen T, Conway BR. Representation of Perceptual Color Space in Macaque Posterior Inferior Temporal Cortex (the V4 Complex). eNeuro. 2016;3.

Brainard DH. Cone contrast and opponent modulation color spaces in *Human Color Vision* (P.K. Kaiser and R.M. Boynton Editors, 2nd Edition). Washington, DC: Optical Society of America. 1996.

Brouwer GJ, Heeger DJ. Decoding and reconstructing color from responses in human visual cortex. J Neurosci. 2009;29:13992-4003.

Carlson T, Tovar DA, Alink A, Kriegeskorte N. Representational dynamics of object vision: the first 1000 ms. J Vis. 2013;13.

Cauchoix M, Barragan-Jason G, Serre T, Barbeau EJ. The neural dynamics of face detection in the wild revealed by MVPA. J Neurosci. 2014;34:846-54.

Cichy RM, Pantazis D, Oliva A. Resolving human object recognition in space and time. Nat Neurosci. 2014;17:455-62.

Cicmil N, Bridge H, Parker AJ, Woolrich MW, Krug K. Localization of MEG human brain responses to retinotopic visual stimuli with contrasting source reconstruction approaches. Front Neurosci. 2014;8:127.

Clifford CW, Spehar B, Solomon SG, Martin PR, Zaidi Q. Interactions between color and luminance in the perception of orientation. J Vis. 2003;3:106-15.

Conway BR. Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). Journal of Neuroscience. 2001;21:2768-83.

Conway BR. Color vision, cones, and color-coding in the cortex. Neuroscientist. 2009;15:274-90.

Conway BR. The Organization and Operation of Inferior Temporal Cortex. Annu Rev Vis Sci. 2018;4:381-402.

De Valois KK, Switkes E. Simultaneous masking interactions between chromatic and luminance gratings. Journal of the Optical Society of America. 1983;73:11-8.

Delahunt PB, Brainard DH. Does human color constancy incorporate the statistical regularity of natural daylight? J Vis. 2004;4:57-81.

Derrington AM, Krauskopf J, Lennie P. Chromatic mechanisms in lateral geniculate nucleus of macaque. Journal of Physiology. 1984;357:241-65.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006;31:968-80.

Dijkstra N, Mostert P, Lange FP, Bosch S, van Gerven MA. Differential temporal dynamics during visual imagery and perception. Elife. 2018;7.

Dobkins K. Moving colors in the lime light. Neuron. 2000;25:15-8.

Dobs K, Isik L, Pantazis D, Kanwisher N. How face perception unfolds over time. Nat Commun. 2019;10:1258.

Ennis RJ, Zaidi Q. Geometrical structure of perceptual color space: Mental representations and adaptation invariance. J Vis. 2019;19:1.

Favreau OE, Cavanagh P. Color and luminance: independent frequency shifts. Science. 1981;212:831-2.

Flachot A, Gegenfurtner KR. Processing of chromatic information in a deep convolutional neural network. J Opt Soc Am A Opt Image Sci Vis. 2018;35:B334-B46.

Gegenfurtner KR. Cortical mechanisms of colour vision. Nat Rev Neurosci. 2003;4:563-72.

Gegenfurtner KR, Rieger J. Sensory and cognitive contributions of color to the recognition of natural scenes. Curr Biol. 2000;10:805-8.

Goddard E, Carlson TA, Dermody N, Woolgar A. Representational dynamics of object recognition: Feedforward and feedback information flows. Neuroimage. 2016;128:385-97.

Goddard E, Chang DHF, Hess RF, Mullen KT. Color contrast adaptation: fMRI fails to predict behavioral adaptation. Neuroimage. 2019;201:116032.

Granger EM, Heurtley JC. Letters to the editor: Visual chromaticity-modulation transfer function. J Opt Soc Am. 1973;63:1173-4.

Haynes JD, Rees G. Decoding mental states from brain activity in humans. Nat Rev Neurosci. 2006;7:523-34.

Hebart MN, Bankson BB, Harel A, Baker CI, Cichy RM. The representational dynamics of task and object processing in humans. Elife. 2018;7.

Horwitz GD, Hass CA. Nonlinear analysis of macaque V1 color tuning reveals cardinal directions for cortical color processing. Nat Neurosci. 2012;15:913-9.

Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology. 1968;195:215-43.

Isik L, Meyers EM, Leibo JZ, Poggio T. The dynamics of invariant object recognition in the human visual system. J Neurophysiol. 2014;111:91-102.

Jin X, Costa RM. Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature. 2010;466:457-62.

Johnson EN, Hawken MJ, Shapley R. Cone inputs in macaque primary visual cortex. J Neurophysiol. 2004;91:2501-14.

Johnson EN, Hawken MJ, Shapley R. The orientation selectivity of color-responsive neurons in macaque V1. J Neurosci. 2008;28:8096-106.

Kaiser D, Azzalini DC, Peelen MV. Shape-independent object category responses revealed by MEG and fMRI decoding. J Neurophysiol. 2016;115:2246-50.

King J-R, Wyart V. The Human Brain encodes a Chronicle of Visual Events at each Instant of Time. bioRxiv. 2019:846576.

King JR, Dehaene S. Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn Sci. 2014;18:203-10.

Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. in Advances in Neural INformation Processing systems. 2012:1097-105.

Kuffler SW. Discharge patterns and functional organization of mammalian retina. J Neurophysiol. 1953;16.

Lafer-Sousa R, Conway BR. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. Nat Neurosci. 2013;16:1870-8.

Lafer-Sousa R, Conway BR. #TheDress: Categorical perception of an ambiguous color image. J Vis. 2017;17:25.

Lafer-Sousa R, Conway BR, Kanwisher NG. Color-Biased Regions of the Ventral Visual Pathway Lie between Face- and Place-Selective Regions in Humans, as in Macaques. J Neurosci. 2016;36:1682-97.

Lafer-Sousa R, Hermann KL, Conway BR. Striking individual differences in color perception uncovered by 'the dress' photograph. Curr Biol. 2015;25:R545-6.

Lafer-Sousa R, Liu YO, Lafer-Sousa L, Wiest MC, Conway BR. Color tuning in alert macaque V1 assessed with fMRI and single-unit recording shows a bias toward daylight colors. J Opt Soc Am A Opt Image Sci Vis. 2012;29:657-70.

Lee BB, Martin PR, Valberg A. Sensitivity of macaque retinal ganglion cells to chromatic and luminance flicker. Journal of Physiology. 1989;414:223-43.

Lennie P, Krauskopf J, Sclar G. Chromatic mechanisms in striate cortex of macaque. Journal of Neuroscience. 1990;10:649-69.

Lindsey DT, Brown AM. Universality of color names. Proc Natl Acad Sci U S A. 2006;103:16608-13.

Logothetis NK, Schiller PH, Charles ER, Hurlbert AC. Perceptual deficits and the activity of the color-opponent and broad-band pathways at isoluminance. Science. 1990;247:214-7.

MacLeod DI, Boynton RM. Chromaticity diagram showing cone excitation by stimuli of equal luminance. J Opt Soc Am. 1979;69:1183-6.

Mannion DJ, McDonald JS, Clifford CW. Discrimination of the local orientation structure of spiral Glass patterns early in human visual cortex. Neuroimage. 2009;46:511-5.

Marti S, Dehaene S. Discrete and continuous mechanisms of temporal selection in rapid visual streams. Nat Commun. 2017;8:1955.

Marti S, King JR, Dehaene S. Time-Resolved Decoding of Two Processing Chains during Dual-Task Interference. Neuron. 2015;88:1297-307.

Martin Cichy R, Khosla A, Pantazis D, Oliva A. Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. Neuroimage. 2017;153:346-58.

Martin PR, White AJ, Goodchild AK, Wilder HD, Sefton AE. Evidence that blue-on cells are part of the third geniculocortical pathway in primates. Eur J Neurosci. 1997;9:1536-41.

Maunsell JH, Ghose GM, Assad JA, McAdams CJ, Boudreau CE, Noerager BD. Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. Vis Neurosci. 1999;16:1-14.

Meyers EM. The neural decoding toolbox. Front Neuroinform. 2013;7:8.

Mullen KT. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. Journal of Physiology. 1985;359:381-400.

Nassi JJ, Callaway EM. Parallel processing strategies of the primate visual system. Nat Rev Neurosci. 2009;10:360-72.

Nealey TA, Maunsell JH. Magnocellular and parvocellular contributions to the responses of neurons in macaque striate cortex. J Neurosci. 1994;14:2069-79.

Pearce B, Crichton S, Mackiewicz M, Finlayson GD, Hurlbert A. Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. PLoS One. 2014;9:e87989.

Quentin R, King JR, Sallard E, Fishman N, Thompson R, Buch ER, et al. Differential Brain Mechanisms of Selection and Maintenance of Information during Working Memory. J Neurosci. 2019;39:3728-40.

Rabin J, Switkes E, Crognale M, Schneck ME, Adams AJ. Visual evoked potentials in three-dimensional color space: correlates of spatio-chromatic processing. Vision Res. 1994;34:2657-71.

Rafegas I, Vanrell M. Color representation in CNNs: parallelisms with biological vision. Ieee Int Conf Comp V. 2017:2697-705.

Reid RC, Shapley RM. Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus. J Neurosci. 2002;22:6158-75.

Reinagel P, Reid RC. Precise firing events are conserved across neurons. J Neurosci. 2002;22:6837-41.

Sandhaeger F, von Nicolai C, Miller EK, Siegel M. Monkey EEG links neuronal color and motion information across species and scales. Elife. 2019;8.

Seymour K, Clifford CW, Logothetis NK, Bartels A. Coding and binding of color and form in visual cortex. Cereb Cortex. 2010;20:1946-54.

Shevell SK, Kingdom FA. Color in complex scenes. Annu Rev Psychol. 2008;59:143-66.

Smet KA, Webster MA, Whitehead LA. A simple principled approach for modeling and understanding uniform color metrics. J Opt Soc Am A Opt Image Sci Vis. 2016;33:A319-31.

Solomon SG, Lennie P. Chromatic gain controls in visual cortical neurons. J Neurosci. 2005;25:4779-92.

Teichmann L, Grootswagers T, Carlson TA, Rich AN. Seeing versus knowing: The temporal dynamics of real and implied colour processing in the human brain. Neuroimage. 2019;200:373-81.

Thorell LG, De Valois RL, Albrecht DG. Spatial mapping of monkey V1 cells with pure color and luminance stimuli. Vision Research. 1984;24:751-69.

Toga AW, Thompson PM, Mori S, Amunts K, Zilles K. Towards multimodal atlases of the human brain. Nat Rev Neurosci. 2006;7:952-66.

Tong F, Pratte MS. Decoding patterns of human brain activity. Annu Rev Psychol. 2012;63:483-509.

van de Nieuwenhuijzen ME, Backus AR, Bahramisharif A, Doeller CF, Jensen O, van Gerven MA. MEG-based decoding of the spatiotemporal dynamics of visual category perception. Neuroimage. 2013;83:1063-73.

van der Horst GJ, Bouman MA. Spatiotemporal chromaticity discrimination. J Opt Soc Am. 1969;59:1482-8.

Wardle SG, Kriegeskorte N, Grootswagers T, Khaligh-Razavi SM, Carlson TA. Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. Neuroimage. 2016;132:59-70.

Westland S, Ripamonti C, Cheung V. Computational colour science using matlab (2nd Edition). John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England. 2012.

Wiesel TN, Hubel DH. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. J Neurophysiol. 1966;29:1115-56.

Winkler AD, Spillmann L, Werner JS, Webster MA. Asymmetries in blue-yellow color perception and in the color of 'the dress'. Curr Biol. 2015;25:R547-8.

Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. Computer Vision - Eccv 2014, Pt I. 2014;8689:818-33.