Temporal dynamics of the neural representation of hue and luminance contrast

Katherine L. Hermann[1*^], Shridhar R. Singh[1*], Isabelle A. Rosenthal[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, and SS analyzed the data; IR, SS and BRC made the figures; DP provided expertise and resources for MEG; BRC supervised the work and wrote the paper.

Acknowledgements: We thank Chris Baker, Susan Wardle, Santani Teng, and Qasim Zaidi for helpful discussions. This research was supported in part by the Intramural Research Program of the National Institutes of Health, National Eye Institute, and grant R01 EY-0233223 to BRC.

Hue and luminance contrast are the most basic visual features, emerging in early layers of convolutional neural networks trained to perform object categorization. In human vision, the timing of the neural computations that extract these features, and the extent to which they are determined by the same or separate neural circuits, is unknown. We addressed these questions using multivariate analyses of human brain responses measured with magnetoencephalography. We report four discoveries. First, it was possible to decode hue tolerant to changes in luminance contrast, and luminance contrast tolerant to changes in hue, consistent with the existence of separable neural mechanisms for these features. Second, the decoding time course for luminance contrast peaked 16-24 ms before hue and showed a more prominent secondary peak corresponding to decoding of stimulus cessation. These results support the idea that the brain uses luminance contrast as an updating signal to separate events within the constant stream of visual information. Third, neural representations of hue generalized to a greater extent across time, providing a neural correlate of the preeminence of hue over luminance contrast in perceptual grouping and memory. Finally, decoding of luminance contrast was more variable across participants for hues associated with daylight (orange and blue) than for anti-daylight (green and pink), suggesting that color-constancy mechanisms reflect individual differences in assumptions about natural lighting.

The most basic computations performed by the visual system yield hue and luminance contrast. *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. Although hue and luminance contrast are often depicted as independent dimensions[1], the relationship between these features is not well understood[2, 3]. On the one hand, the demonstration of independent frequency shifts for color and luminance[4] and the selective impact of adaptation on detection thresholds[5, 6] suggests that hue and luminance contrast are processed by separate neural channels. Separate encoding is also evident in convolutional neural networks trained for object recognition, which show independent filters for hue and luminance contrast in the earliest layer[7-10]; this feature of CNNs may reflect the independence of color and luminance edges in natural scene statistics[11]. Hue and luminance contrast are also distinguished by their efficiency in visual memory tasks: hue is more efficient than luminance contrast[12]. On the other hand, the many perceptual interactions of hue and luminance contrast suggest these features are processed together. These interactions are evident in the perception of orientation[13], judgments of brightness across colors[14], color categorization[15], and color naming. For example, across cultures, languages link warm colors (e.g. *orange*) with high luminance contrast (*light*), and cool colors (e.g. *blue*) with low luminance contrast (*dark*) [16]; such associations may reflect joint statistics in natural lights (sunlight is warm and bright, versus shadow, which is cool and dark). Interactions of hue and luminance contrast are also evident in masking. Hue can mask changes in luminance contrast[17]. Consider the 8 spirals in Figure 2a; most people group them by hue (rows), not luminance contrast (columns). Taken together, the behavioral data underscore long-standing questions about the extent to which hue and luminance contrast are encoded by the same or separate neural mechanisms. Also unresolved is a fundamental question about the amount of time that the brain takes to extract these features.

One approach to identifying neural mechanisms involves interrogating the feed-forward flow of visual information from the photoreceptors to cortex. Individual parvocellular neurons of the lateral geniculate nucleus, which reflect the first post-receptoral stage of visual processing, respond to both hue and luminance contrast[18, 19], consistent with combined encoding of these dimensions. Furthermore, parvocellular cell responses correspond to the relative spatial frequency of luminance-contrast vision (high-pass) and color vision (low-pass)[17, 20-22]. But the LGN also contains magnocellular cells, which are sensitive only to luminance contrast, and koniocellular neurons, some of which respond to S-cone signals and may

represent a distinct chromatic channel[23, 24]. These observations animate the alternative idea, that the LGN encodes hue and luminance contrast in separate channels[25]. The LGN data are also inconclusive about the amount of time that the brain takes to compute hue and luminance contrast. If parvocellular neurons encode both hue and luminance contrast, these computations should take the same amount of time. Alternatively, if these dimensions 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[26].

Color must depend on mechanisms in visual cortex, where neurons become sensitive to mixtures of the subcortical channels. A prominent theory is that within visual cortex, 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[27]. This theory is supported by the observation that most V1 cells receive a mixture of inputs from subcortical channels and respond to both color and luminance contrast[27-31]. But a small population of V1 cells show striking responses to color stimuli that lack luminance contrast[30, 32-34]. Thus the V1 neurophysiology cannot rule out the possibility that hue and luminance contrast could be encoded by parallel pathways[35].

Clues to the neural mechanisms of color and luminance contrast are provided by univariate visual evoked potential measurements to equiluminant and achromatic stimuli[36, 37]. But it has not been possible to infer from these experiments the underlying neural mechanisms because all subcortical channels respond to equiluminant stimuli[38]. 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[39]. Clues have also been provided by fMRI[40-42], but fMRI does not predict results from behavioral adaptation[6], and cannot uncover the relevant temporal dynamics because it is limited by the relatively sluggish time course of blood flow[43] .

Here we explore a new approach to address the neural mechanisms for encoding hue and luminance contrast, using magnetoencephalography (MEG) in humans, coupled with multivariate analysis[44-48] (**Figure 1**). Color can be decoded from MEG activity[49-56]. In the present work, we exploit the exquisite temporal resolution of MEG to tease apart the neural mechanisms for hue and luminance contrast.

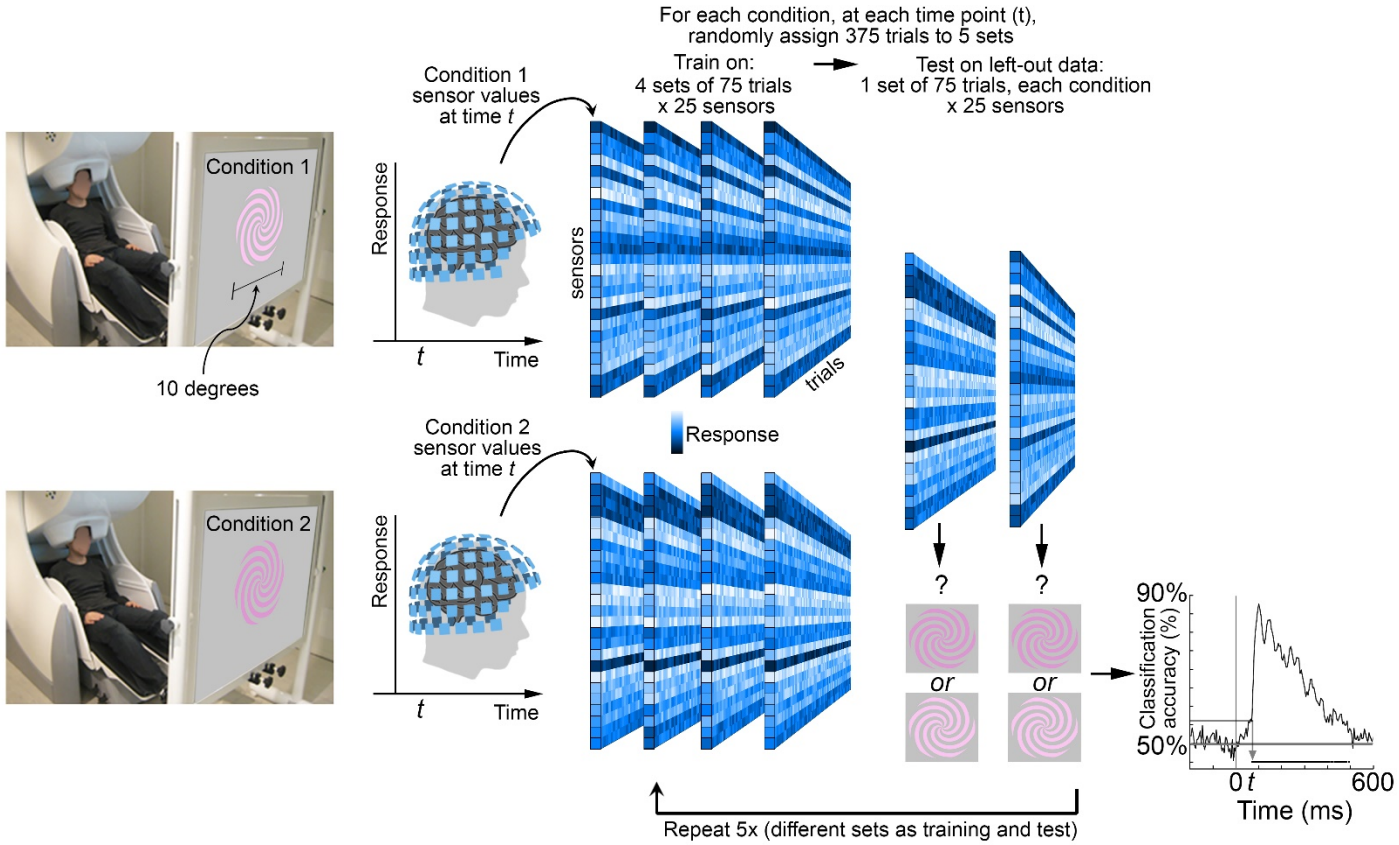


Figure 1. Experimental paradigm

Participants were scanned with MEG while they looked at colored spirals that were flashed on the screen for 116 ms (with 1 second between flashes). The spirals were one of eight colors: four hues at two luminance contrast levels (**SI Figure 1** provides the color specifications for the stimuli; see also Figure 2 a,c). Stimuli were pseudo-randomly interleaved, with 500 presentations of each stimulus over two recording sessions for each of 18 participants. 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, 200ms before stimulus onset to 600ms after stimulus onset. The figure shows simulated data to illustrate the analysis pipeline. Classifiers were trained at every time point, and individually for each participant: 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, yielding a decoding curve showing classification accuracy as a function of time after stimulus onset (bottom right showing actual data for classifying luminance contrast using pink). The horizontal sequence of data points above the x-axis shows the time points when decoding was significant for more than 5 consecutive time points (FDR corrected; the gray arrow shows the classification accuracy and the time point at which decoding became significant).

Results

The experiments were designed to enable a decoding analysis of MEG responses, to answer two over-arching questions. First, given patterns of MEG activity, is it possible to decode the luminance contrast and hue of the stimulus that elicited the activity? Second, what is the time course – that is, how long does it

take the brain to extract these features? We measured MEG responses in 18 participants while they were shown spirals that could appear in one of eight colors (**Figure 2a**). The colors were defined by the DKL color space, which is constructed from the cone-opponent dimensions that correspond to retinal color-encoding mechanisms (**SI Figure 1**) [57, 58]. The 8 colors consisted of four hues at two luminance levels (4 x 2 design). The stimuli were set to the intermediate directions in DKL color space; thus all stimuli involved modulations of both L-M cone activity and S-cone activity, and all stimuli had the same magnitude of modulation of L-M, and the same magnitude of modulation of S. Different colors were created by pairing different signs of these modulations. The luminance contrast of all the stimuli was fixed (26%) but varied in sign (light or dark) relative to the 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. In other words, the results should show evidence that hue decoding generalizes across luminance contrast; and luminance-contrast decoding generalizes across hue. 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 invariant to the other dimension. Given these alternatives, it was important for the experiment to have enough power, which we ensured by first conducting an extensive pilot experiment to determine the number of trials needed to obtain reliable data (**SI Figure 2**).

We used a maximum correlation coefficient classifier (as implemented in the Neural Decoding Toolbox [59], see Methods). We performed within-participant decoding (trained and evaluated classifiers on data in each participant), independently for each timepoint (applied independent classifiers at each time point relative to stimulus onset). All analyses were cross validated (**Figure 1**), yielding plots that show how representations unfold over time. Participants were told to maintain fixation throughout stimulus presentation and to blink at designated times. 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 in hue. The impact of task was assessed in a control

experiment, in which two participants performed half the trials using a 1-back hue matching task and the other half of the trials using a 1-back luminance-contrast matching task. Task had no impact on the main conclusions, as described below (see **Figure 4;** see **SI Figure 5** for behavioral results). We used a spiral-shaped stimulus to avoid cardinal or radial response biases [60-62].

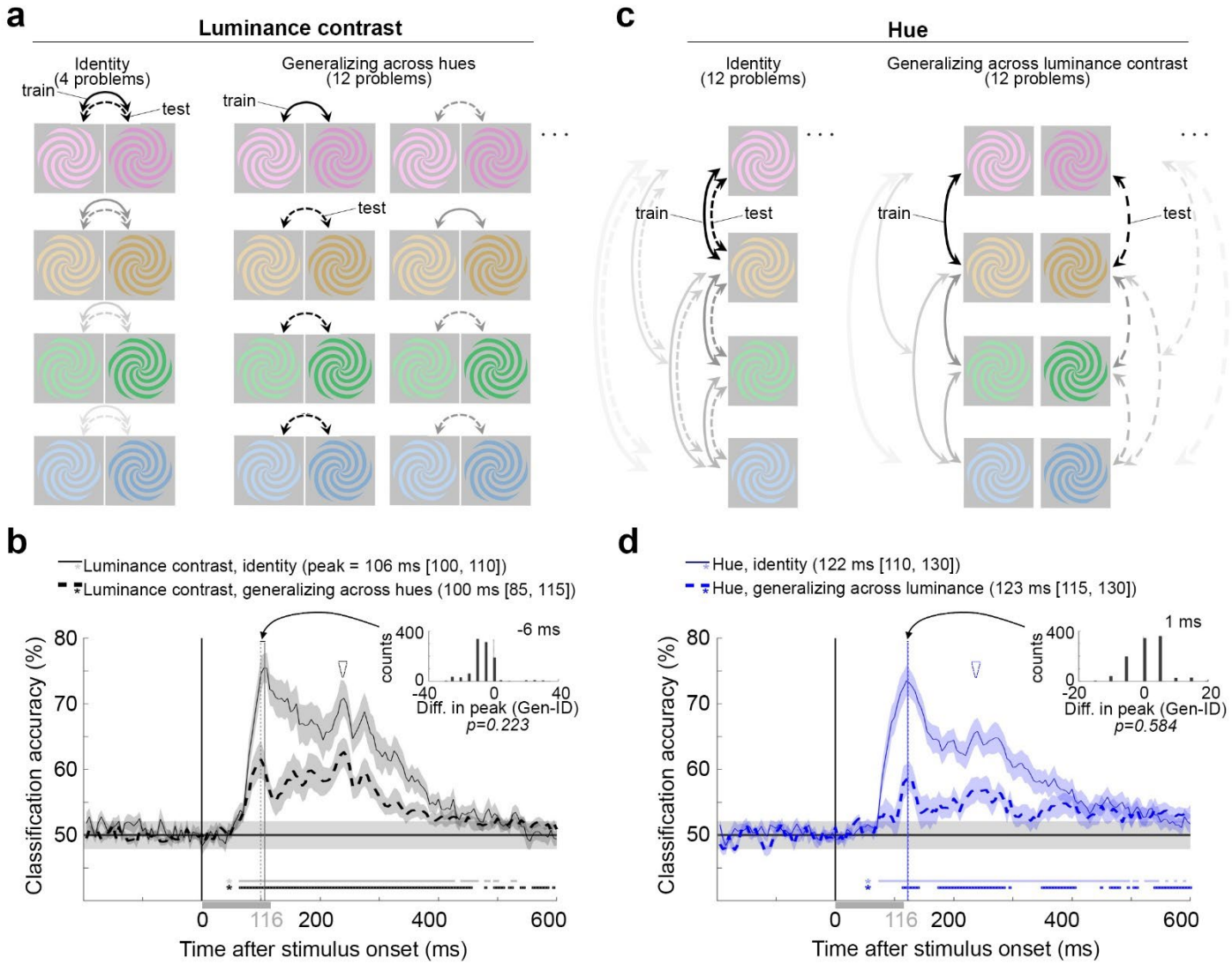


Figure 2. Decoding luminance contrast and hue from MEG data

a) Decoding luminance contrast. 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 generalizing-across-hue 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 (all four identity problems are illustrated in the graphic). For the generalization problems, classifiers were trained and tested on different hues (half of the 12 generalizing-across-hue problems are illustrated). In the graphic, solid lines indicate comparisons used for training and dashed lines indicate comparisons used for testing; line shading distinguishes different problems. **b)** Classification accuracy as a function of time after stimulus onset for the luminance contrast problems (the average of the four identity problems, solid line; the average of the 12 generalization problems, dashed line). The traces were generated by

averaging the 1000 bootstrapped samples across the 18 participants. Shading around the decoding curves shows the standard error of the bootstrap samples; for reference, shading around the line at 50% shows the 95% CI of the decoding performance prior to stimulus onset (chance); the stimulus duration was 116 ms (gray bar along the x axis). The inset shows the difference in peak for the 1000 bootstrapped comparisons: identity minus generalization (mean = -6 ms, p=0.223). Initial peak of the identity problems, solid vertical gray line, 106 ms [100 110]; initial peak of the generalization problems, dashed gray line, 99 ms [85 115]. Open arrowhead shows the second peak, which corresponds to stimulus cessation. The horizontal sequence of data points above the x-axis, demarcated by asterisks, show the time points at which decoding was above chance for at least five consecutive time bins (FDR corrected); onset of significant decoding for the identity problems was achieved at 68 ms [50, 75], for the generalization problems, at 62 ms [50, 80]. **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 the luminance contrast was opposite (e.g. light), requiring generalization of hue across luminance contrast (12 generalizing-across-luminance 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 = 1 ms, p=0.584) for the 1000 bootstrapped comparisons across participants. The time to peak was the same for the identity and generalization problems (identity problems: solid vertical line, 122 ms [110, 130]; generalization problems: dashed vertical line, 123 ms [115, 130]). Onset of significant decoding for the identity problems was achieved at 71 ms [35, 80], and for the generalization problems, at 119 [20, 365]. Other conventions as for panel (b).

The eight colors ("conditions"; see **Figure 2a**) were presented in pseudo-random order for 116 ms with 1 second of the gray background between presentations (1 sec ISI). 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 1** 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;** the stimuli associated with the training of classifiers are indicated by solid arrows, and the stimuli associated with the corresponding tests are shown by dashed arrows). In other problems of luminance-contrast decoding, the hue 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 problems as generalizing-across-hue since they uncover the extent to which luminance contrast can be decoded independent of hue (**Figure 2a, right**). In other analyses we determined the extent to which classifiers could

decode hue identity (**Figure 2c, left**), and hue generalizing-across-luminance-contrast (**Figure 2c, right**). The generalization problems provide a test of invariance: decoding luminance-contrast invariant to hue, and decoding hue invariant to luminance contrast. **SI Figure 3** shows the test-retest reliability of the data from the main experiments [63]. All analyses involved binary classifiers (chance is 50%) to facilitate a direct comparison of the results among the different problems.

There is no accepted metric for equating color contrast and luminance contrast [39, 64]. The lack of an accepted metric has made it hard to compare responses to equiluminant stimuli and luminance-contrast stimuli using univariate analyses because response magnitude will depend on stimulus contrast [36]. Multivariate analyses overcome this hurdle because they uncover any difference in the pattern of response between a pair of variables, providing a general framework that provides insight about the information processed by the brain [65]. This is advantageous in the context of the present work because any evidence of significant decoding of hue generalizing across luminance contrast, and luminance contrast generalizing across hue, would not depend on relative contrast of luminance and color, so long as the stimuli are suprathreshold. The stimuli were well above detection threshold and were approximately matched in units of detection threshold (see Methods).

Decoding luminance contrast

Figure 2b shows the average classifier performance across the four luminance-contrast identity problems (solid line) and the 12 generalizing-across-hue problems (dashed line; shading around each trace shows the bootstrapped standard error). The results show that classification accuracy was significantly above chance for luminance contrast for both types of problems. Thus it was possible to decode luminance contrast independent of hue, which supports the hypothesis that the brain has a representation of luminance contrast that is independent of the representation of hue. The magnitude of decoding does not provide a measure of the absolute size of an effect, but it is nonetheless a valid measure of relative effect sizes within a given study [65]. Peak decoding accuracy was higher for the luminance-contrast identity problems compared to the luminance-contrast generalization problems (76% [71, 80] versus (62% [57, 67]; square brackets contain the 95% C.I, which was obtained by bootstrapping.). This result supports the hypothesis that the brain also has a representation of luminance contrast that is combined with the representation of

hue. In other words, these results show that the brain has a representation of luminance contrast carried by channels that are hue selective.

Let us compare the time course of decoding for the identity problems and the generalization problems. The latency at which decoding became significant was not different for the identity problems (68ms [50, 75]) versus the generalization problems (62 ms [50, 80]) based on bootstrapping (n=18, p=0.36). The time to peak was not different for the identity problems compared to the generalization problems (inset Figure 2b, p=0.223; dashed vertical line, 100 ms [85, 115], versus solid vertical line, 106 ms [100, 110]; confidence limits computed by 1000 bootstrap draws across participants). Close inspection of the data shows a trend in favor of an earlier time to peak for the generalization problems (generalization problems showed equal or earlier time to peak, p=0.035); this observation implies that the time-to-peak is a reliable measure of how much time the brain takes to generate a representation, because it does not vary in a trivial way with changes in the amplitude of decoding (one might have expected time-to-peak to be longer for curves with lower amplitude decoding; the data show, if anything, the opposite: the problem with lower amplitude decoding is slightly earlier). Following peak decoding, the generalization problems showed a pronounced dip. Finally, both the identity and generalization problems had a second prominent decoding peak (open arrowhead), following the initial peak (curved arrow). 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

Figure 2d shows the average performance across the 12 hue identity problems (solid line) and the 12 generalizing-across-luminance-contrast problems (dashed line; other conventions as in Figure 2b). The plot shows that hue was decodable in both cases. That it was possible to decode hue generalizing across luminance supports the hypothesis that the brain has a representation of hue that is separate from the representation of luminance contrast. The latency at which decoding became significant was 119 ms [20, 365] for the generalization problems, and 71 ms [35, 80] for the identity problems (the latency of the generalization problems was greater than the upper 95% CI limit for the latency of the identity problems, but a direct comparison of the bootstrapped values was not significant, p=0.19). As with the luminance-contrast decoding problems, decoding had a higher peak magnitude for the identity problems (74% [69,

78]) compared to the generalization problems (59% [55, 64]), which provides 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 (solid vertical line, 122 ms [110, 130]; dashed vertical line, 123 ms [115, 130]; p=0.584). Moreover, the time to peak was not significantly different for identity problems among high luminance contrast stimuli (124 ms [105, 140]) compared to identity problems among low luminance contrast stimuli (121 ms [110, 130]); and the magnitude of peak decoding accuracy for these two sets of problems also were not significantly different (for light stimuli: 71% [67, 76]; for dark stimuli: 76% [71, 81]) (**SI Figure 4**). These results show that any potential difference in saturation between the light and dark stimuli does not influence the time course of decoding (saturation is ill-defined, but as discussed in **SI Figure 1**, there is an argument that the saturation of the dark stimuli is higher than the saturation of the light stimuli). Finally, to the extent the solid line in **Figure 2d** shows a second peak (open arrowhead, corresponding to decoding stimulus cessation), 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).

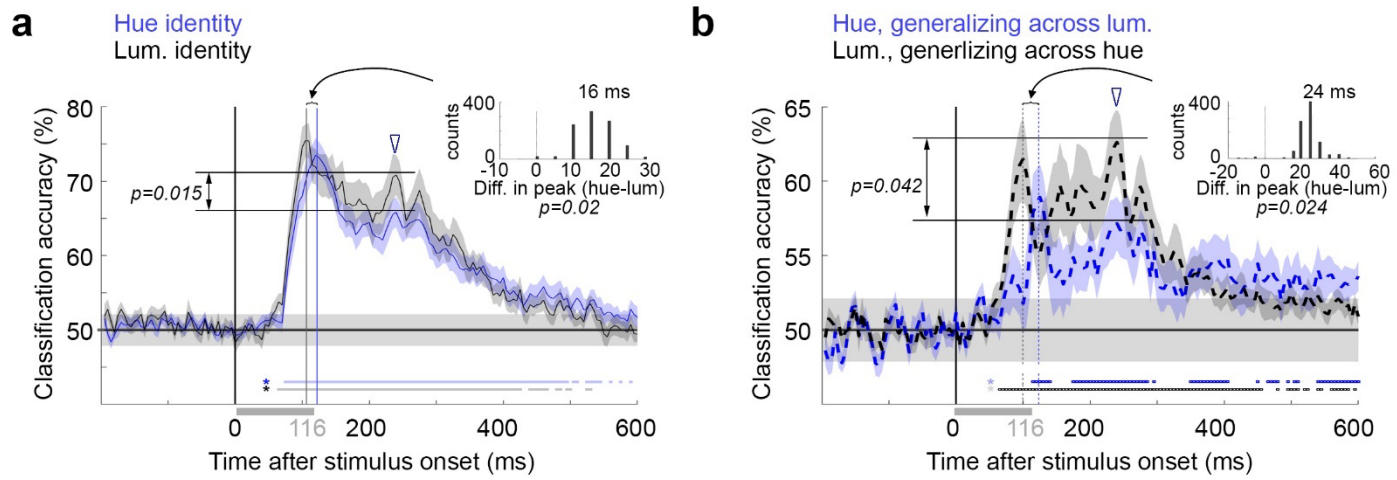


Figure 3. Comparing the temporal dynamics of decoding, luminance contrast versus hue
a) Classification accuracy for the identity problems (see Figure 2a,c). Inset shows the differences between the peaks across the 1000 bootstrapped samples across the 18 participants. Accuracy for luminance contrast peaked 16 ms before hue (p=0.02). The magnitude of classifier accuracy corresponding to the cessation of the stimulus (open arrowhead) was greater for luminance contrast than hue (p=0.015). **b)** Classification accuracy for the generalization problems. Accuracy for luminance contrast peaked 24 ms before hue (p=0.024). The magnitude of classifier accuracy corresponding to the cessation of the stimulus was greater for luminance contrast than hue (p=0.042). Other conventions as Figure 2 b,d.

A comparison of the hue and luminance-contrast identity problems (**Figure 3a**) and of the hue and luminance-contrast generalization problems (**Figure 3b**) underscores three differences in time course of decoding. First, hue was decodable after luminance contrast as assessed by the time of peak decoding (identity problems: 16 ms delay (p=0.02); generalization problems: 24 ms delay (p=0.024). Differences in the latency of decoding onset showed a trend that reflects these timing differences, but the differences in latency of decoding onset were not significant (for identity problems, p=0.528; for generalization problems, p=0.09). Second, the time point of peak hue decoding corresponded to the dip in the luminance decoding curve, and vice versa, the time point of peak luminance decoding corresponded to a notch in the hue decoding curve. This is especially evident in the plots of the generalization problems. This was especially clear in the generalization curves, where the black and blue traces are in counterphase (quantified as the Spearman correlation of the derivative of the decoding curves, computed for 116 ms following the average of the onset latencies of the hue and luminance generalization problems; R=-0.55, p=0.007). Third, the second peak (corresponding to decoding stimulus cessation) was larger for luminance contrast than for hue (double-headed arrow; for the identity problems, p=0.015; for the generalization problems, p=0.042). By comparison, the magnitude of the initial peaks was not different (for the identity problems, p=0.214; for the generalization problems, p=0.238).

Impact of task on decoding

Decoding accuracy of visual stimuli from patterns of MEG activity could be impacted by task engagement during MEG data collection [66, 67]. Task effects can enhance task-relevant object features and are reported to arise relatively late after stimulus onset [68]. This literature raises the question: should the differences in timing for decoding hue versus luminance contrast be attributed to the task that participants performed during MEG data collection? We answered this question in a control experiment conducted prior to the main experiments. The control experiment was identical to the main experiments except that for half the runs the participants performed a 1-back hue-matching task, and for the other half of the runs the participants performed a 1-back luminance-contrast matching task (**Figure 4a;** participants completed 5 sessions; each session had 20 runs; each run had 100 trials; each trial involved a 116 ms stimulus presentation and a 1 sec ISI; the task alternated between runs in each session). Participants learned both tasks, achieving excellent accuracy within the first MEG data-acquisition session and improving in reaction time over consecutive sessions (**SI Figure 5**).

The time to peak for decoding hue was different than the time to peak for decoding luminance contrast, for data collected using both tasks (**Figure 4a**; difference in time to peak for data obtained using the 1-back luminance task, p=0.004; difference in time to peak for data obtained using the 1-back hue task, p<0.001; note that precise p values cannot be reported for values smaller than p=0.001 because the p values were derived from 1000 bootstrap iterations). The time to peak for decoding hue was 110 ms [110, 110] using data obtained with the hue task, and 110 ms [110, 115] using data obtained with the luminance task; while the time to peak for decoding luminance contrast was 90 ms [85, 95] (hue task), and 89 ms [85, 90] (luminance task). These results replicate the results from the main experiments, showing that decoding luminance contrast reached a peak earlier than decoding hue. The time to peak decoding of luminance contrast was indistinguishable for data collected under the two task conditions (**Figure 4b, left panel**; p=0.384); similarly the time to peak decoding of hue was indistinguishable for data collected under the two task conditions (**Figure 4b, right panel**; p=0.427). These results show that task did not impact the time to peak, for decoding hue or luminance contrast. The data collected in the control experiment also show that the relative timing difference for decoding hue versus luminance contrast was apparent for each individual subject: combining data from both tasks, the time to peak for decoding hue for subject 1 was 110 ms [100, 110], and for subject 2 was 110 ms [110, 110]); the time to peak for decoding luminance contrast for subject 1 was 90 ms [85, 90], and for subject 2 was 90 [75, 95]). The time to peak for decoding hue was different from the time to peak for decoding luminance contrast in both subjects (both S1 and S2, p=0.001). In addition, the data from the control experiment also confirm that the secondary peak in the decoding curve was more prominent for the luminance-contrast problems compared to the hue problems.

The results of the control experiment provide evidence of one possible impact of task: the duration of significant decoding. Decoding of luminance was significant for a longer duration when participants were engaged in the 1-back luminance task compared to when they were engaged in the 1-back hue task (**Figure 4b, left panel**, horizontal lines below the decoding trace show time points of significant decoding). Similarly, decoding of hue was significant for a longer duration when participants were engaged in the 1-back hue task compared to when they were engaged in the 1-back luminance task (**Figure 4b, right panel**). The results of the control experiment show that task did not have an impact on the time course of decoding, except at relatively late time points (>300 ms after stimulus onset), which is consistent with the manifestation of task-related effects documented by others [68].

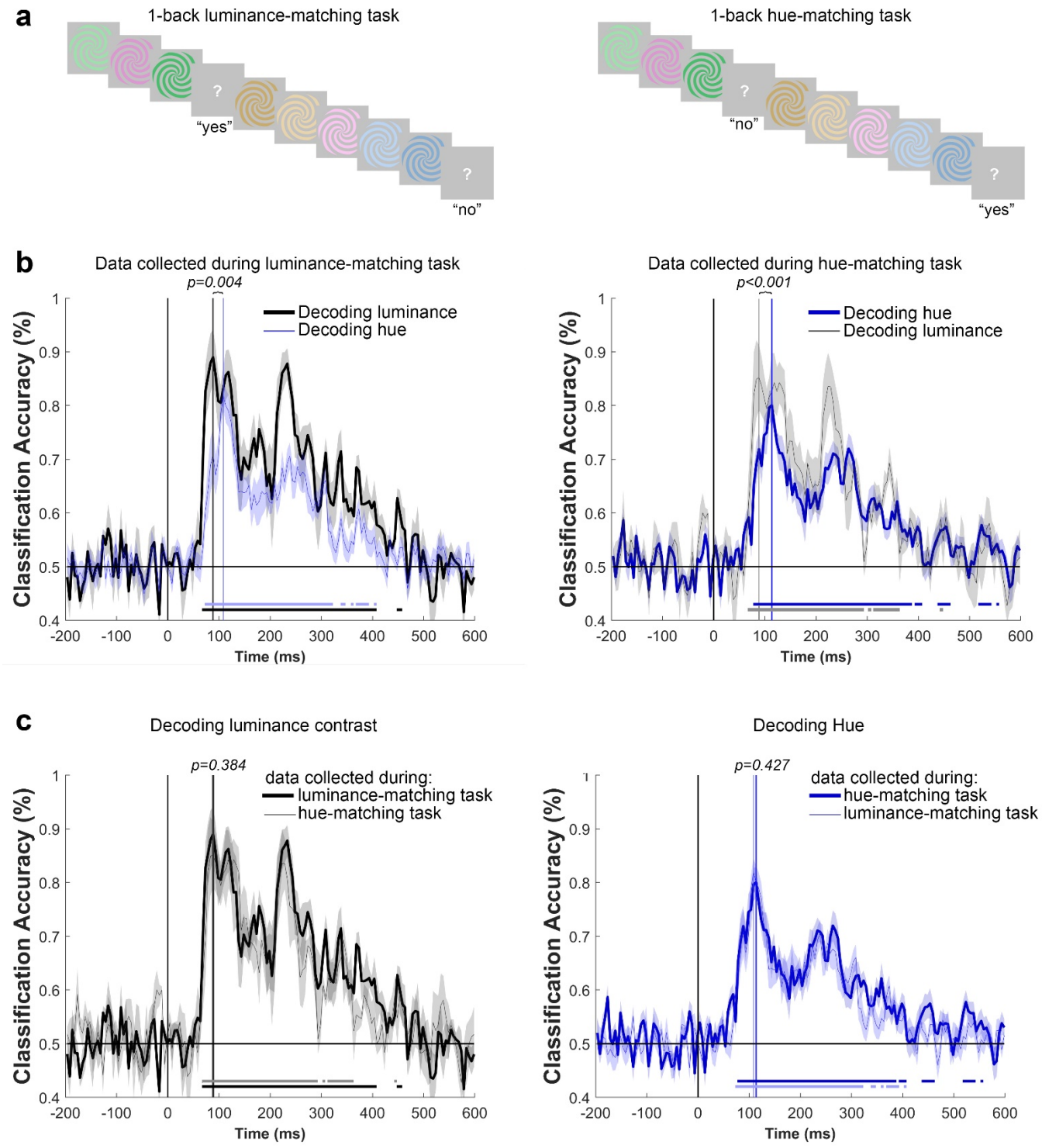


Figure 4. Task had no impact on the time to peak decoding of hue or luminance contrast. a) left panel, illustration of the 1-back luminance-contrast matching task. Every 3-5 stimulus presentations, participants were shown a "?" and asked to report via button press whether the two preceding stimuli matched in luminance contrast or not. The correct answers are provided below the "?". **Right panel,** illustration using the same stimulus sequence as the left panel, showing the 1-back hue-matching task. **b) left panel,** Decoding luminance contrast (black trace) and hue (blue trace) from data collected while participants performed the 1-back luminance-contrast matching task. The data show the average of the identity decoding problems illustrated in Figure 2a,c. The vertical lines show the time to peak for

the two traces. The horizontal sequence of points above the x-axis show time points when decoding was significantly greater than chance for five or more consecutive time bins (significance tests on individual time bins were FDR corrected). Hue decoding peaked after luminance-contrast decoding (p=0.004). **Right panel,** Decoding luminance contrast (black line) and hue (blue line) of data collected while participants performed the 1-back hue matching task. Hue decoding peaked after luminance-contrast decoding (p<0.001). Other conventions as for the left panel. Analyzing data from each subject separately, the peak in decoding was later for hue compared to luminance (p=0.001 for both subjects), and the time to peak for decoding hue was different from the time to peak for decoding luminance contrast (both S1 and S2, p=0.001). **c), left panel,** decoding luminance contrast using data collected while participants performed the 1-back luminance-contrast matching task (bold line), or the 1-back hue matching task (thin line); the time-to-peak for decoding luminance contrast identity was not impacted by task (p=0.384; the vertical lines showing the time to peak for the two sets of problems are at the same time point). **Right panel,** decoding hue using data collected while participants performed the 1-back hue matching task (bold line), or the 1-back luminance-contrast matching task (thin line); the time-to-peak for decoding hue was not impacted by task (p=0.427). The bolded traces show the decoding problems that aligned with the task. Luminance-contrast decoding was significant for a longer duration for data collected during the luminance-matching task (left panel), while hue decoding was significant for a longer duration for data collected during the hue-matching task (right panel).

Greater cross-temporal decoding for hue than luminance contrast

The results discussed so far evaluate the classifiers' performance using test data obtained at the same time point after stimulus onset as the data used to train the classifier. The classifiers show significant decoding for a substantial amount of time, tens or hundreds of milliseconds, following stimulus onset. One possibility is that the pattern of activity is relativity stable over this time period; another possibility is that it is dynamic [69]. To distinguish between these alternatives, we trained classifiers on the patterns of activity at each point in time and evaluated the extent to which each could predict activity at all other time points. If activity patterns are dynamic, the cross-temporal 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.

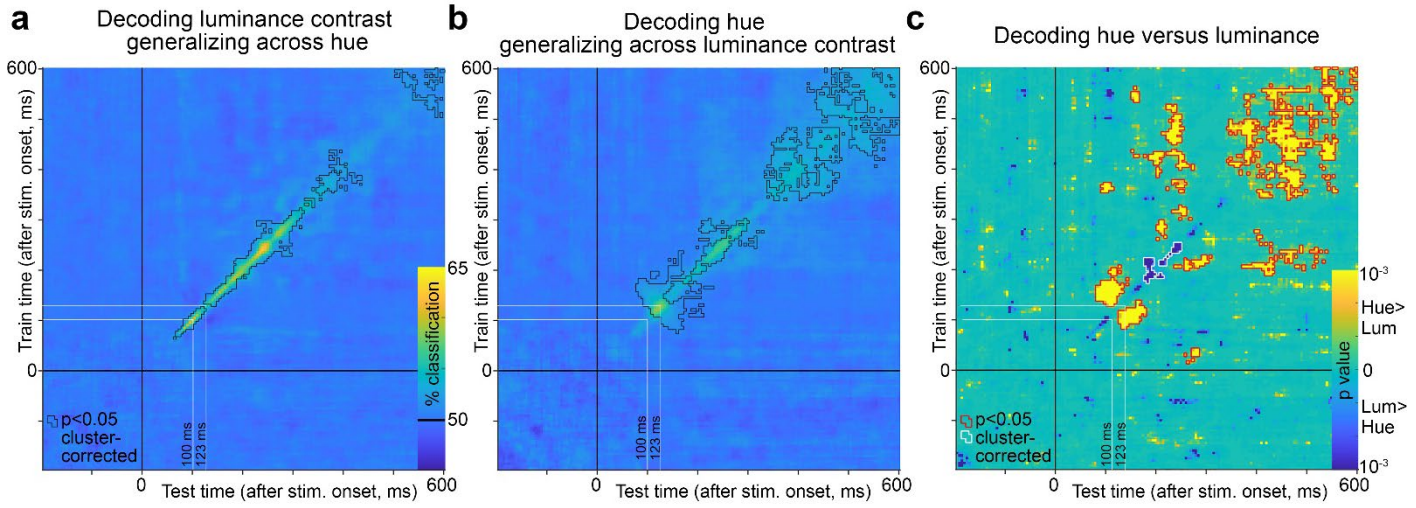


Figure 5. Cross-temporal generalization of decoding, luminance contrast versus hue

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 for decoding luminance contrast generalizing 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 123 ms. The black contours show regions in the heatmap that were p<0.05 cluster corrected. **b**) Data as in panel (a), but for classifiers trained and tested on the set of problems for decoding hue generalizing across luminance contrast. The peak classification time was 123 ms; the cluster-corrected regions of significance extended further from the diagonal compared to panel (a). **c**) Comparison of results in panel (a) and (b). The time points where the classifiers were more accurate for luminance contrast compared to hue are shown as dark blue (white contours show cluster-corrected significant results), while the time points where the classifiers were more accurate for hue compared to luminance contrast are shown as yellow (red contours show cluster-corrected results). The p values were obtained by bootstrapping over problems and FDR corrected. Decoding of hue showed greater generalization across time compared to decoding of luminance contrast, and the greater cross-temporal generalization for hue began relatively early after stimulus onset (~123 ms).

Figure 5 shows the cross-temporal plots for decoding luminance contrast generalizing across hue (**Figure 5a, left**), and for decoding hue generalizing across luminance contrast (**Figure 5a, right**); the results show the results averaged over the 12 individual problems for luminance contrast and hue. 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 black contours identify data that were deemed significant by a bootstrap test and 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 5a** compared to **Figure 5b**. **Figure 5c** 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 relatively early after stimulus onset, emerging at about the same time as the initial peak in hue generalization decoding (123ms).

These results support the hypothesis that the patterns of activity in the brain associated with hue 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) [70-74]. The chromaticity of natural lights is linked to relative luminance contrast: direct sunlight has a warm spectrum (e.g. orange), and has relatively high luminance contrast, whereas indirect light such as in shadows, has a cool spectrum (reflecting skylight) and a relatively low luminance contrast. If the visual system is adapted to natural statistics, we predicted that luminance contrast carried by orange and blue would be less meaningful about object boundaries than luminance contrast carried by green and pink, because luminance contrasts linked to orange and blue chromaticity are more likely to be attributed to the illumination. Testing the extent to which luminance contrast can be carried by different hues therefore has important implications for understanding the neural mechanisms that support color constancy [70, 71, 75].

Individual luminance-contrast decoding problems averaged over participants showed substantial variability both in the four identity problems (**Figure 6a**) and the twelve generalization problems (**Figure 6b**). **Figure 6c** shows luminance-contrast decoding performance for each problem during the 5 ms time bin set by the average peak decoding for the luminance-contrast identity problems (105ms-110ms). Data along the inverse diagonal correspond to the identity problems (recall the identity problems are those in which the hue was the same for training data and testing data). Data off the diagonal correspond to decoding problems for luminance contrast generalizing across hue, and these data allow us to test the hypothesis. The 95% CI of the classification performance is shown for each problem, obtained by bootstrapping. The significance of decoding was determined by a permutation test (see Methods). Luminance contrast could not be decoded above chance when the train and test stimuli were blue and orange, but it could be decoded when the train and test stimuli were pink and green, providing support for the hypothesis.

Perhaps luminance contrast could be decoded at a different time point than the time point of peak decoding for the average of the luminance-contrast identity problems? To address this possibility, **Figure 6d** shows the peak decoding performance at the time of peak decoding determined separately for each of the sixteen luminance-contrast problems. The time bin of peak decoding for each problem 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. The only two problems that did not achieve significant decoding performance for more than 5 consecutive bins were those problems in which the train and test stimuli were blue and orange.

The data in **Figures 6 c,d** support three findings. First, luminance-contrast decoding was strongest when the carrier (training) hue was the same as the test hue (unpaired two sided t test on average peak decoding values comparing on-diagonal to off-diagonal **Figure 6c**, p=0.0005; **Figure 6d**, p=0.0014). Second, among these identity problems, classification performance was higher for warm colors (pink, orange) than cool colors (blue, green; paired two sided t test across individual average warm color performance and individual average cool color performance, **Figure 6c**, p=0.007; **Figure 6d**, p=4e-5, supporting our prior analysis [54]). 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 6c**, p=0.02; **Figure 6d**, p=0.03.

The statistical analysis shown in Figure 6 was obtained by comparing and averaging results across participants. But the time course of decoding accuracy may show individual differences that will not be apparent in these averages. Thus a final possibility we consider is that the lack of significant luminance-contrast decoding for blue and orange in Figure 6 reflects variability in the timing and amplitude of decoding across participants, rather than a lack of success in decoding luminance contrast from orange and blue within each participant. To address this possibility, we first determined the average decoding accuracy throughout the decoding time course (0-600ms) for each problem, for each participant (72 problems total: 18 participants x 2 daylight-axis problems x 2 anti-daylight-axis problems). Luminance contrast could be decoded using anti-daylight colors (54.1% [51.9, 56.2], and using daylight colors (53% [50.8, 54.5]), with no difference (p=0.236). Next, for each problem, we calculated the Spearman correlation of the decoding time course for each of the 18 participants with the time course of another participant drawn at random from the remaining 17 participants, to estimate the temporal correlation of the decoding curves. We repeated this 1000 times and averaged across the repetitions to obtain a mean estimate of the temporal correlation of each participant's time course to those of the others. If the time course of decoding shows the same structure across participants, the analysis will yield high temporal correlations. We found that the temporal correlation was, on average, higher for problems involving anti-daylight colors versus daylight colors (p=0.0435). Problems involving daylight colors were different from problems involving anti-daylight colors in both their mean accuracy and temporal correlation (MANOVA, p=0.0007), confirming that the temporal correlation of luminance contrast decoding problems is more variable across participants for orange and blue than for pink and green. (Note that the correlation of the average decoding accuracy with the temporal correlation (**SI Figure 6**) is not surprising, because when decoding is higher, the temporal correlation will be more apparent.) These analyses support the hypothesis that luminance contrast can be decoded from orange and blue, but that the time course and magnitude of decoding luminance contrast is more impacted by individual differences for these problems compared to decoding luminance contrast from pink and green.

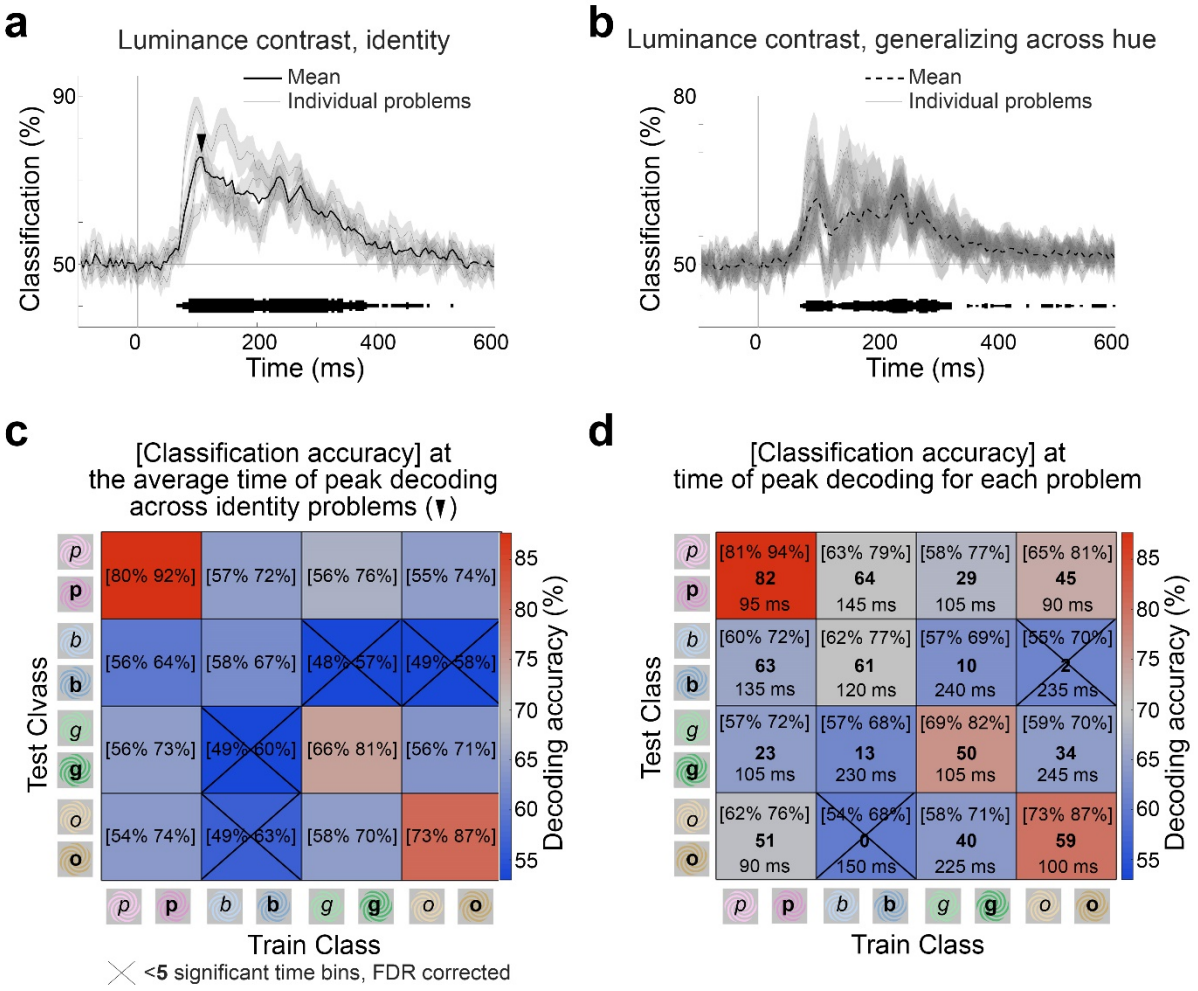


Figure 6. Decoding luminance contrast carried by different colors

a) Accuracy of classifiers trained to decode luminance contrast for stimuli of each hue (4 identity problems, thin lines; average, thick line). **b)** Accuracy of classifiers trained to decode luminance contrast across hues (12 generalizing-across-hue problems, thin lines; average, thick dashed line; shading shows standard error across the 18 subjects). For panels (a) and (b), the thickness of horizontal line at the bottom of each graph shows the number of decoding problems (of 4, left; of 12, right) that were significant in each 5ms bin (FDR corrected). **c)** Heatmap showing the 95% CI of the classifier performance for each problem at the average time of peak decoding across the luminance-identity problems (105-110 ms after stimulus onset, arrowhead in panel a). Subproblems that were not significant for more than five consecutive bins are marked with an X. **d)** Heatmap showing the peak decoding accuracy for each of the 16 binary classifiers at the time point of peak decoding for each problem**.** The bold numbers show the longest stretch of consecutive time bins with significant classification performance for each problem. The time stamps show when the classifier performance reached peak decoding accuracy. Other conventions as for panel c. Luminance contrast was not successfully decoded when training on responses to blue (or orange) and testing on response to orange (or blue); but it was successfully decoded when training and testing using pink and green.

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 [76]. 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 7a 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 [77]. By aligning each participant to a standard atlas [78, 79] 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 7b**, 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 two sided t-test); but were different from LO (p=0.01), and from PPA (p=0.02) (**Figure 7c**). 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 7d**). Hue generalized across luminance contrast was not decodable in any region except to a very small extent in V2 (**Figure 7e**). The distribution of sensors used in the decoding analysis is shown in **Figure 7f, g** (see legend and methods for details).

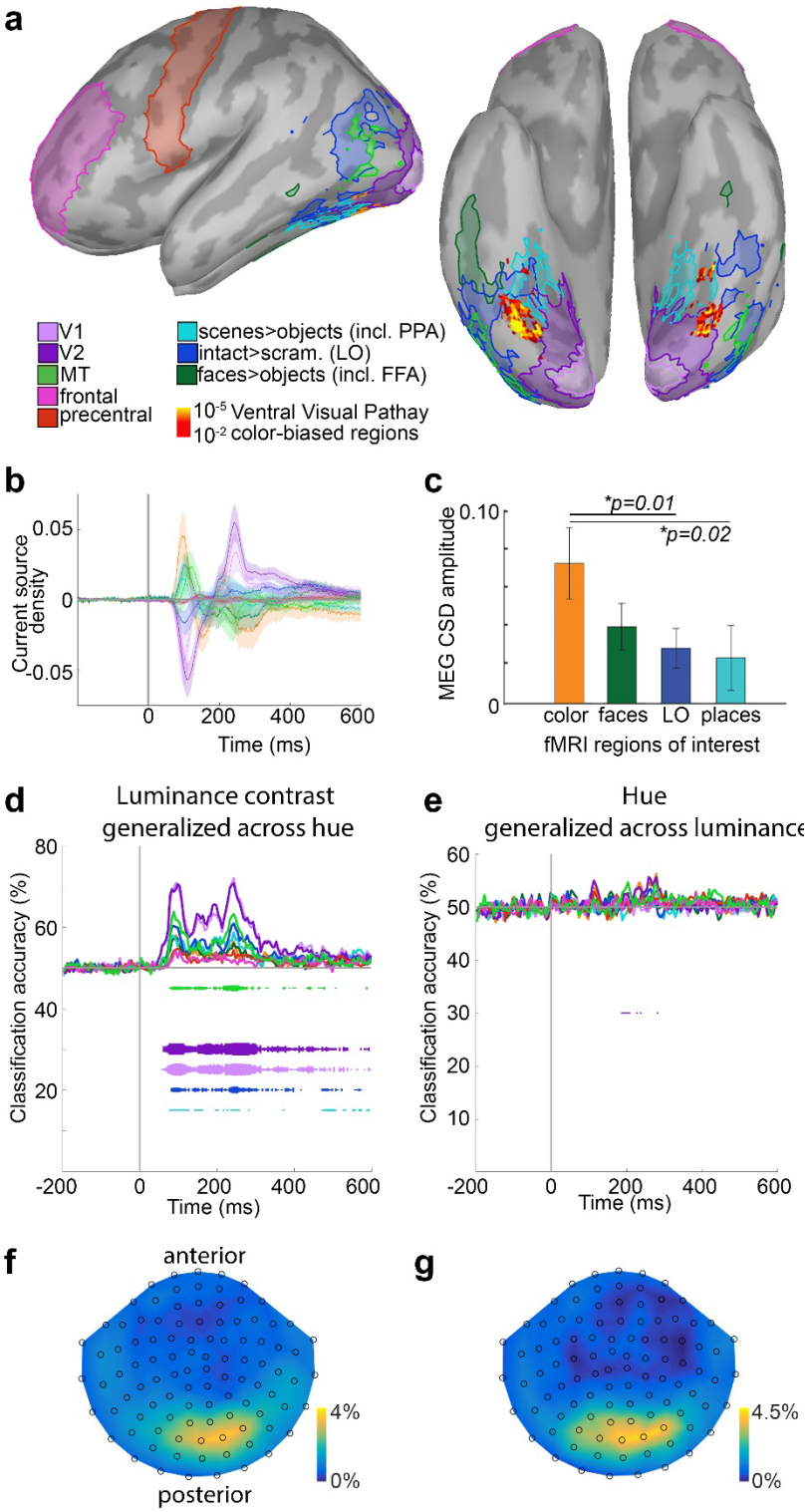


Figure 7. Source localization to regions defined using fMRI

a) Example functional anatomy from one participant shown on the inflated cortical surface (left panel shows side view, anterior to the left; right panel shows ventral view, anterior at top). Regions of interest were defined using fMRI responses to 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 (yellow-orange) evident in the ventral view 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 using an anatomical atlas. **b)** Source localization estimates of the current source density of MEG responses to color arising from each functional ROI, 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)** Maximum amplitude of CSD in each ROI of the ventral visual pathway, 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 decoding luminance contrast generalizing-across-hue (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 decoding hue generalizing-across-luminance contrast (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 colored stimuli and produced five main findings: first, hue and luminance contrast could be decoded independently. Luminance contrast could be decoded even if the training and testing stimuli differed in hue; and hue could be decoded even if the training and testing stimuli differed in luminance contrast. Second, the time course of decoding was different for hue compared to luminance contrast (**Figure 3**). The peak for decoding hue was 16-24 ms after the peak for decoding luminance contrast; and the decoding time course for luminance contrast showed a prominent second peak corresponding to stimulus cessation. Control experiments showed that these timing differences could not be attributed to the task that participants performed during MEG data collection (**Figure 4**). Third, classification accuracy was higher for problems in which the stimuli used to train and test the classifiers were identical (identity problems, **Figure 2**), compared to problems that required generalization across hue or luminance contrast (generalizing problems, **Figure 2**). Fourth, representations of hue showed greater cross-temporal generalization than representations of luminance contrast. And fifth, representations of luminance contrast varied depending on the hue carrying the luminance-contrast information. Luminance contrast attached to orange and blue (colors associated with the daylight locus) were decoded with more variability across participants than luminance contrast attached to pink and green (colors associated with the anti-daylight locus). Together, these results have implications for our understanding of the different roles that luminance contrast and hue play in visual perception, and how encoding these features is implemented in the neural circuitry.

Multivariate analyses of signals acquired across the brain provide a powerful tool to uncover the way in which perceptual experiences are encoded[80, 81]. And multivariate analyses specifically of MEG data uncover important information about the time taken by the brain to perform computations[44, 45, 51]. In the present work, evidence for significant decoding of hue generalizing across luminance contrast, and luminance contrast generalizing across hue, shows that either these features are encoded to some extent by different neural populations, or if they 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 necessarily multiplexed simultaneously by the same neural population. The results provide, to our knowledge, the first direct neural evidence that the human brain has independent representations for hue and luminance contrast, and the timing of these representations, which has implications for our understanding of how luminance contrast and hue are used in vision.

In object perception, the decoding time course reflects the perceptual and categorical dissimilarity of stimuli, with more perceptually dissimilar stimuli (i.e. higher levels of abstraction) decodable later, and associated with computations performed by areas further along the visual-processing hierarchy [82 44 46] (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. The time course for decoding hue (peak 123 ms; **Figure 3**) is comparable to that for decoding shape-independent object category[83] and face identity[84], operations that probably reflect activity in area LO and the fusiform face area (FFA). The decoding time course for hue is therefore consistent with the hypothesis 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, part of the V4 Complex[77, 85-87]. 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[88-90], showing that these neurons are capable of representing hue tolerant to changes in luminance contrast.

Neural representations related to object vision that emerge earliest, as determined by classifiers tested using the same images on which they were trained (i.e. not requiring any perceptual dissimilarity), reflect the encoding of "low-level visual features"[44]—decoding in these classifiers peaks early (<100ms), and is attributed to operations implemented early in the visual-processing hierarchy, perhaps V1 [83, 91-93]. But it has not been clear what constitutes the low-level visual features that determine the early decoding performance. It is often implied that the low-level features consist of both oriented luminance-contrast edges and color, an implication reinforced by the observation of separate luminance-contrast edge filters and color filters in the earliest layer of convolutional neural networks[7]. But the early time to peak of luminance-contrast decoding (100 ms), which was 16-24 ms earlier than peak hue decoding, suggests that (1) luminance contrast (and not hue) is the first feature to be encoded; and (2) luminance contrast is implemented at an early stage in the visual-processing hierarchy, before hue. The relative timing difference for decoding hue and luminance contrast may be relatively small—just tens of milliseconds—but it is not only robust, evident within individual subjects (**Figure 4**), but also long enough to encompass 7-10 synapses given monosynaptic transmission latencies of ~2.5 ms in the visual pathway[94]. The results underscore the

exquisite temporal precision of MEG and leverage it to answer a question about the timing of fundamental operations of the visual system that have, to date, been intractable.

Knowledge of the timing of the neural events is important, in part because it provides clues to the different roles that 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 hundreds of milliseconds[66]. Presumably the visual system must rapidly parse this stream of information to enable encoding of new content. But it must do so while retaining 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[69]. 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[95]. In object vision, dynamical systems modeling predicts the existence of observable update mechanisms that signal new content[96]. 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 (**Figure 5a**). The time course for decoding hue, meanwhile, showed a clear peak only at stimulus onset and showed relatively stronger cross-temporal generalization (**Figure 5b**). 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 [97]—and it is 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.

What, then, does the more stable representation of hue, compared to luminance contrast, suggest about the role of hue in vision? People with normal color vision find it very difficult to make heterochromatic brightness matches [17], 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). These behavioral observations show that, under normal circumstances, hue is a more salient dimension than luminance contrast, and one that is more sustained: across glances, humans are more likely

to retain visual memory for hue than for luminance contrast[12] (the challenge of making heterochromatic luminance-contrast matches is evident in the behavioral results in the present work, see **SI Figure 5**). The present results provide a plausible neural correlate for the relatively more sustained representation of hue, and they may help resolve a long-standing paradox in visual neurophysiology. In the earliest studies of V1 single-unit responses, Hubel and Wiesel remarked that "it was surprising to us, however, that the great majority of cells [in V1] could discriminate precisely the orientation or direction of movement of a stimulus, but had no marked selectivity regarding wave-length"[98]. Their surprise derived from the fact that hue is such a prominent dimension in perception. The present results suggest that this prominence is achieved not by marked selectivity to color, but by more stable patterns of response to color across the population. 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 could therefore be better exploited for grouping and remembering visual information.

While the results show that representations of hue and luminance contrast are somewhat decoupled, they also indicate the existence of representations that combine hue and luminance contrast: 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 therefore consistent with the idea that in addition to separate representations of hue and luminance contrast, the visual system has representations that combine hue and luminance contrast, perhaps reflecting the joint selectivity for hue and luminance contrast evident in some neural populations in the LGN[18] and V1[27, 34, 99]. These combined representations may provide a neural correlate for the interactions of hue and luminance contrast evident in perception, as reviewed in the introduction.

Decoding of luminance contrast varied depending on which hue was carrying the luminance-contrast signal. Across participants, 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. Classifiers trained and tested using pink and green, meanwhile, could

successfully decode luminance contrast (**Figure 6**). This asymmetry derives not because luminance contrast cannot be decoded from orange and blue, but rather because luminance-contrast decoding using orange and blue had a more variable time course across participants (**SI Figure 6**). We investigated asymmetries in the representation of luminance contrast because we were interested in knowing the extent to which the visual system is adapted to natural lighting conditions. Under natural conditions, the chromaticity of the illuminant covaries with relative luminance contrast, ranging from warm and bright (e.g. direct sunlight) to cool and dark (e.g. shadow). Luminance contrast signals associated with orange and blue chromaticities are therefore less reliable indicators of object boundaries since they may derive from boundaries introduced by cast shadows—this is what is thought to account for the striking individual differences in color appearance of the famous #thedress image[72-74, 100]. Indeed, the choices people make for what constitutes "achromatic" vary most substantially along the daylight locus[70, 71]. The MEG results imply that the neural representation of luminance contrast is adapted to natural lighting conditions and provide a neural substrate for the variability in achromatic settings among people.

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. 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[77]; this pattern is also seen in macaque monkeys[85]. 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[101]. Source-localized analyses did not recover significant luminance-invariant hue representations, which may not be surprising given fundamental limitations of source localization[102].

Methods

Visual Stimuli

Stimuli were eight square-wave spiral gratings on a neutral gray background (**Figure 2a**) [60-62]. The 8 stimulus colors, four hues at two luminance-contrast levels, of matched absolute cone contrast, were defined in DKL color space [12, 57, 58] using implementations by Westland [103] and Brainard [104]: 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). The absolute L-M cone modulation of the stimuli were matched; and the absolute S cone modulation of the stimuli were matched. Different colors were created by pairing different signs of modulation along L-M and S. 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%, computed as Weber contrast because the stimuli were brief and not full field. Modulation of the cone-opponent mechanisms, shown in **SI Figure 1**, was computed relative to the adapting background gray, Judd corrected, using the Smith and Pokorny cone fundamentals [105]. Using MATLAB scripts of Westland et al. [103], the xyY values of the stimuli were converted to LMS values using *xy2MB*, and the LMS values were converted to DKL values using *lms2dkl*. The stimuli were well above detection threshold. In luminance and hue, the stimuli had comparable contrast in units of detection threshold, where luminance contrast was computed in units of detection threshold for luminance; and hue contrast was computed in units of detection threshold for color contrast. The absolute cone contrast ($|L| + |M|$) required to reach detection threshold is about 9x higher for luminance compared to color [12]. Using the values provided by Sachtler and Zaidi, the luminance contrast of the light and dark pink stimuli were ~37x luminance detection threshold, where $|L| = |(L_{light\_pink} - L_{dark\_pink})/ (L_{light\_pink} + L_{dark\_pink})|$, L_{light_pink} is L cone activity elicited by the light pink stimulus and L_{dark_pink} is the L cone activity elicited by the dark pink stimulus; and where $|M| = |(M_{light\_pink} - M_{dark\_pink})/ (M_{light\_pink} + M_{dark\_pink})|$, M_{light_pink} is the M cone activity elicited by the light pink stimulus and M_{dark_pink} is the M cone activity elicited by the dark pink stimulus. The light pink and light green were ~30x above color detection threshold, and the dark pink and dark green stimuli were ~51 above color-detection threshold. By this measure, the contrast of the dark stimuli is higher than the contrast of the light stimuli. This had no impact on the time course of decoding (see **SI Figure 3**).

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). The experimental paradigm was created using psychtoolbox [106]; 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, spoke English as a first language, and provided informed consent. 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 hue-matching 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.

During the MEG data collection, 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. The fixation point 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 (pseudo-randomly 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. Individual runs were identical across subjects, but the order of runs was randomized between subjects. The sequence of stimuli within each run was random with the constraint that the total number of presentations was the same for each stimulus condition over the set of runs obtained for each participant.

In the main experiments, 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.

Pilot experiment to determine the number of trials per condition

Prior to conducing the main experiments, we collected data in four participants (3 female), using four colored stimuli (two hues, blue and orange, at light and dark contrast levels). Participants were instructed to fixate a small spot at the center of the display; besides this passive fixation, there was no task. Eye movements were monitored to ensure passive fixation. The goal of this pilot was to determine the number of trials needed to successfully decode color. As in the main experiment, the stimuli were spirals, subjects viewed each color 500 times, and each stimulus appeared for 100 ms followed by a 1 s ISI. Other details of experimental paradigm was the same as for the main experiment, except that the stimuli included only two hues at two luminance contrast values (four conditions total). The two hues were DKL angles 150 and 300, which are intermediate colors corresponding roughly to blue and yellow; the luminance of the stimuli were: background gray, 41 cd/m^2; high luminance-contrast stimuli, 48-50 cd/m^2; and low luminance-contrast stimuli, 30-32 cd/m^2). To evaluate the experimental power, we set out to present each stimulus with what we thought was an excessive number of trials: 500 (leaving 375 after rejecting trials with eye blinks or other artifacts). **SI Figure 2** shows the data reliability for the pilot experiment. The figure was generated by subsampling independent pairs of N% of data, computing the decoding for each independent set of data, and computing the correlation coefficient between the sets of data at each time point in the decoding curve. As expected, the test-retest curves plateau; but the plateau occurs only with substantial numbers of trials. Informed by these pilot experiments, we performed the main experiment with 500 trials. Preliminary results of these experiments have been presented and provide to our knowledge the first evidence for decoding color from MEG data [49, 50]. The four participants who took part in this pilot experiment did not participate in the main experiments or the control experiment.

Control experiment to measure the impact of task

Before launching the main experiments, we deployed a control experiment to determine the decoding parameters for the main experiments and to evaluate the task effects on decoding. This experiment was conducted in 2 participants (1 female, age 20-30 years) who completed 5 sessions of 20 runs each. Each run had 100 stimulus presentations. Every 3-5 stimulus presentations, participants saw a "?" that prompted a

response regarding the preceding two stimuli; the correct answer depended on the task for the run. For one half of the runs, the participants performed a 1-back hue matching task, and during the other half of the runs they performed a 1-back luminance-contrast matching task (**Figure 4a**). Participants performed the same task for trials in a run, with the task alternating between runs for a given session. The results from the control experiment showed no impact of task performance on the latency or time-to-peak of decoding (**Figure 4 b,c**). For consistency in the main experiment, we had participants perform the same task. The decoding analysis shown in Figure 4b,c, is for the identity problems illustrated in Figure 2a,c. The two participants who took part in the control experiment did not participate in the main experiments or the other pilot experiment.

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) [59]. 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 [47]; 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 for each unique problem. The decoding time courses for each problem were sampled 18 times with replacement and averaged, and this procedure was repeated 1000 times to produce 1000 timecourses, which were averaged to generate the decoding traces in Figures 2 and 3. The gray shading shows the standard error of the bootstrap means.

The 95% CI around the time to peak was determined from the times to peak for each of the 1000

bootstrapped time courses. Statistical tests on the difference in time to peak between two problems were performed using the bootstrap distributions of the differences in time to peak values. If the 95% confidence interval did not include 0, we rejected the null hypothesis, and p values were calculated based on the proportion of values that did fall below 0. To determine at which time points decoding accuracy was significantly above chance, a permutation test was used to calculate p values (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. 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.

The test-retest curves to evaluate experimental power (**SI Figures 2 and 3**) were 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 and computed the temporal correlation between the two decoding time courses. 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. The significance tests were cluster corrected, in which the cluster-size threshold was defined as the largest contiguous area $p<0.05$ over regions of the heat map corresponding to before stimulus onset.

In Figure 5, 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

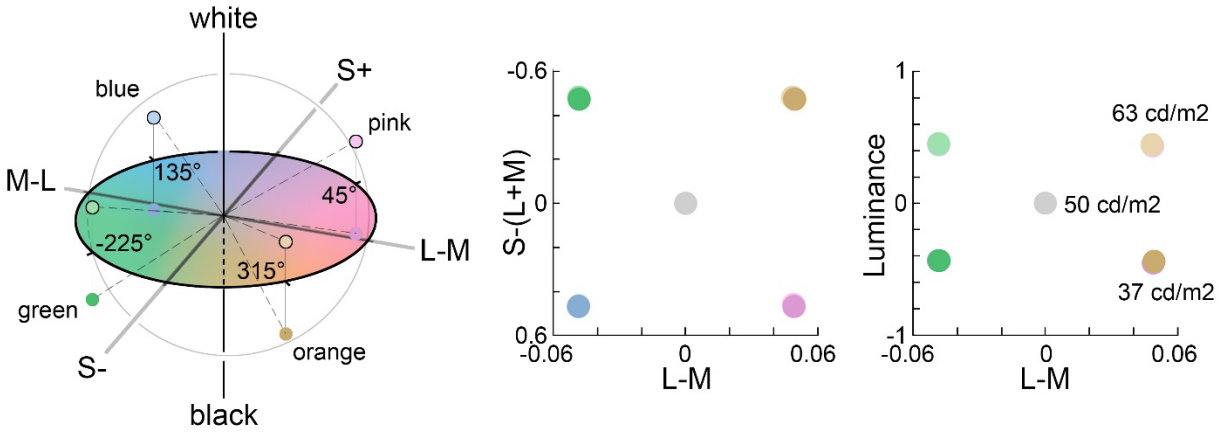
Current source density is a metric representing the current at each point on the surface of the brain, defined by the source grid. First, using Brainstorm, a minimum norm estimate (MNE) was calculated, which was "depth-weighted", to compensate for a bias in current density calculations that results in 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.

Data Availability

All raw datasets analyzed in the current study will be available in the NEICOMMONS repository.

Code Availability

The procedures and code used to produce all the figures, and all statistical analyses, in the current study will be available in the NEICOMMONS repository.



SI Figure 1. Stimulus specification. The colors of the stimuli were set to the intermediate directions of a color space defined by the cardinal cone-opponent mechanisms: the right panel shows the eight colors at hue angles of 45°, 135°, 225°, and 315° within the DKL color space [57, 58]. The right panels show the cone and luminance contrast of the stimuli, relative to the gray adapting background. The xyY values for the stimuli, in the order light/dark pink, blue, green, and orange, were:

0.305229324, 0.293448787, 0.628205196
0.304095045, 0.264040315, 0.370334398
0.272115611, 0.307026806, 0.630601697
0.25265696, 0.284531124, 0.375699736
0.310023683, 0.426308239, 0.63358717
0.312628272, 0.502281154, 0.375716373
0.352384707, 0.400239965, 0.631544192
0.392438648, 0.444792187, 0.374272604

The xyY coordinates of the background gray, Judd corrected, were:
xyY_gray = [0.306634746, 0.346610046, 0.502314308]

The LMS coordinates of the stimuli, computed using the MATLAB function *xy2MB* from Westland [103], which uses the Smith and Pokorny[107] cone fundamentals, were (last row is the background gray):

0.4144	0.2138	0.0138
0.2474	0.1229	0.0097
0.4008	0.2298	0.0139
0.2357	0.1400	0.0098
0.4027	0.2308	0.0063
0.2358	0.1399	0.0022
0.4165	0.2151	0.0063
0.2500	0.1243	0.0022
0.3253	0.1770	0.0081

The DKL values (in radians) of the stimuli (as plotted in the figure above), obtained using the function *lms2dkl* from Westland [103], were (by definition, gray is [0, 0, 0]):

-0.4341	0.0487	-0.4590
0.4551	0.0490	-0.4681
-0.4424	-0.0483	-0.4647
0.4366	-0.0485	-0.4681
-0.4526	-0.0484	0.4815
0.4365	-0.0481	0.4725
-0.4456	0.0484	0.4805

0.4415 0.0493 0.4725

We designed the experiment such that the stimuli were well above detection threshold, and so that the absolute cone contrast of stimuli that differ in hue was comparable to the absolute cone contrast of stimuli that differ in luminance contrast. We can estimate the contrast of the stimuli in units of detection threshold, using detection data provided by Sachtler and Zaidi (1992) [12]. It is important to do this separately for color and luminance contrast because, as Sachtler and Zaidi showed, detection thresholds (measured in absolute cone contrast) are over 9x higher for luminance contrast than for color. The cone contrast is computed as the absolute value of the contrast in L, plus the absolute value of the contrast in M. For example:

We compute the **luminance contrast** carried by the pink hue as follows:

$$|(Lcone_{light\_pink} - Lcone_{dark\_pink})/(Lcone_{light\_pink} + Lcone_{dark\_pink})| + |(Mcone_{light\_pink} - Mcone_{dark\_pink})/(Mcone_{light\_pink} + Mcone_{dark\_pink})|$$

$$= |(0.4144-0.2474)/(0.4144+0.2474)| + |(0.2138-0.1229)/(0.2138+0.1229)|$$

$$= 0.2525 + 0.270 = 0.5225$$

The detection threshold (|L|+|M|) for luminance contrast is 0.014 (from Sachtler and Zaidi [12]). Thus the contrast of the stimuli was (0.5225/0.014), **~37x** detection threshold.

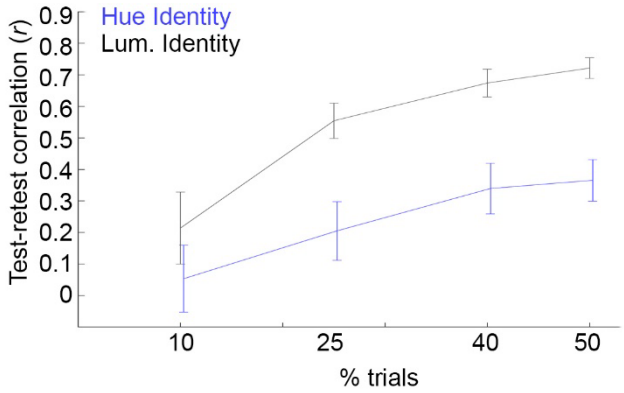
We can compute the **color contrast** between pink and green at a given luminance contrast as follows:

$$|(Lcone_{light\_pink} - Lcone_{light\_green})/(Lcone_{light\_pink} + Lcone_{light\_green})| + |(Mcone_{light\_pink} - Mcone_{light\_green})/(Mcone_{light\_pink} + Mcone_{light\_green})|$$

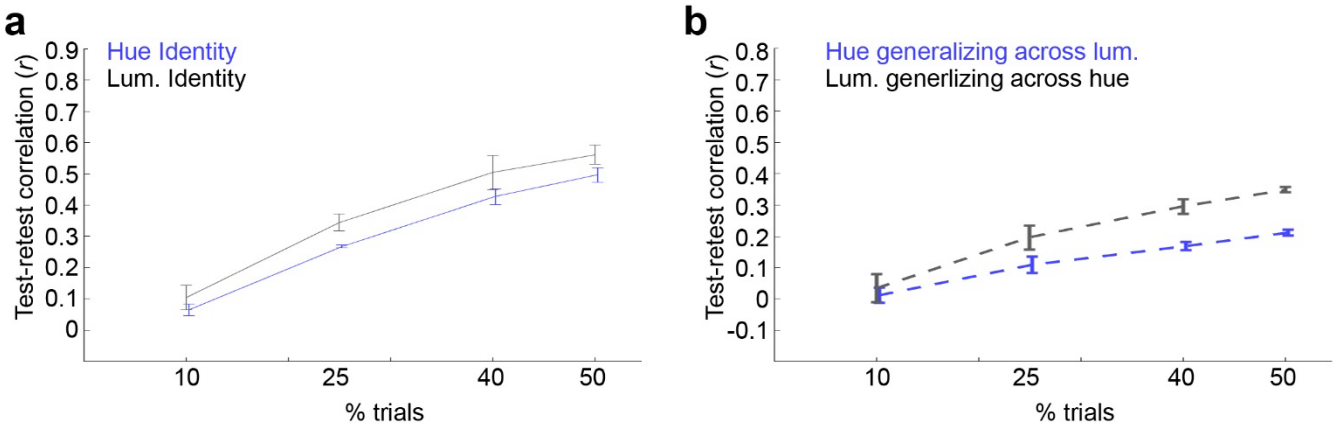
$$= |(0.4144-0.4027)/(0.4144+0.4027)| + |(0.2138-0.2308)/(0.2138+0.2308)|$$

$$= 0.0143 + 0.0382 = 0.0525$$

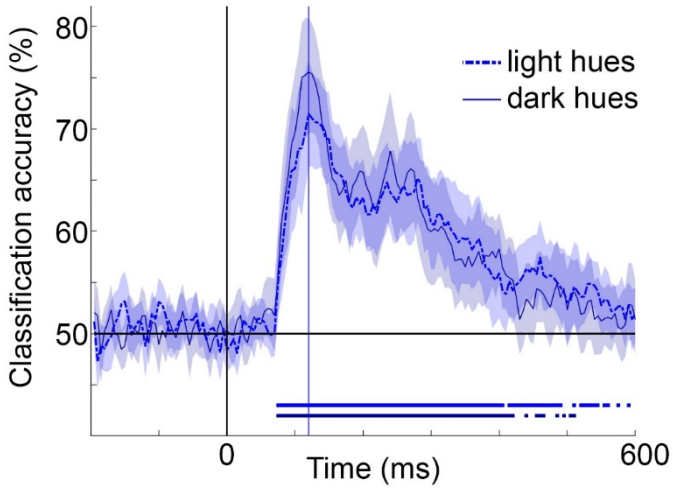
The detection threshold (|L|+|M|) for color contrast is 0.001724 (from Sachtler and Zaidi [12]). Thus the contrast of the light pink and green stimuli we used is (0.0525/0.001724), **~30 x** detection threshold. Comparable calculations for the dark pink and dark green stimuli show these stimuli to be **~51X** detection threshold. The two values of color contrast (for light stimuli and dark stimuli) straddle the absolute contrast (in detection units) across luminance (37x, as calculated above). The possible difference in color contrast of light versus dark stimuli had no impact on the time course for decoding hue (see SI Figure 3), showing that at the supra-threshold values used, decoding strength is not likely influenced by stimulus saturation.



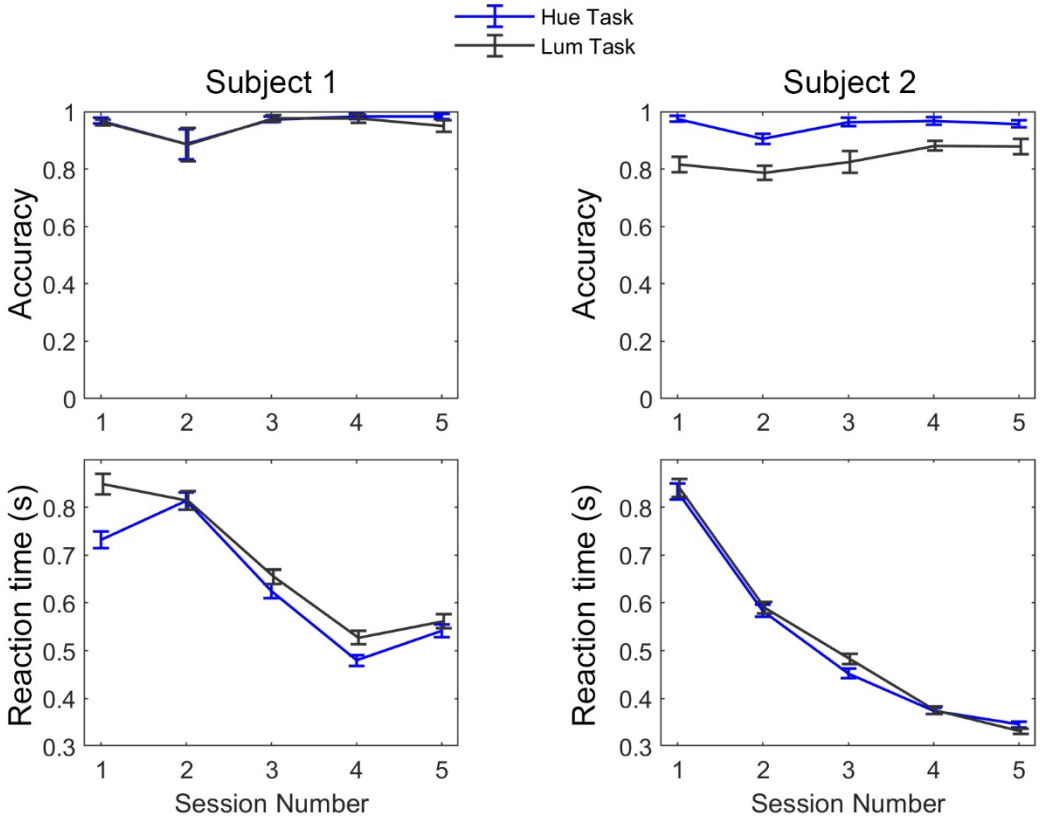
SI Figure 2. Data reliability analysis for the pilot experiment. The graph shows the correlation between pairs of independent data sets, as a function of the amount of data in the data set, for decoding hue identity and luminance-contrast identity. This analysis provides a measure of data reliability [63]. The y-axis shows the correlation coefficient comparing the decoding magnitude at each point in the decoding time course between the two data sets. The graph shows: (1) that the data are reliable (the test-retest estimates are above chance); (2) that the experiments have sufficient power to extract close to as much signal as is possible given the experimental conditions (the test-retest curves come close to plateau when extrapolated to 100% of the data); and (3) that to decode color, one needs relatively many trials. As in the main experiment, the stimuli were spirals, participants viewed each color 500 times, and each stimulus appeared for 100 ms followed by a 1 s ISI. Other details of experimental paradigm were the same as for the main experiment, with the following exceptions: the data were collected in 4 participants (3 female; separate people from those who participated in the main experiments); and the stimuli only included four colors at two luminance contrast values. The two hues were DKL angles 150 and 300, which are intermediate colors corresponding roughly to blue and yellow; the luminance of the stimuli were: background gray, 41 cd/m^2; high luminance-contrast stimuli, 48-50 cd/m^2; and low luminance-contrast stimuli, 30-32 cd/m^2).



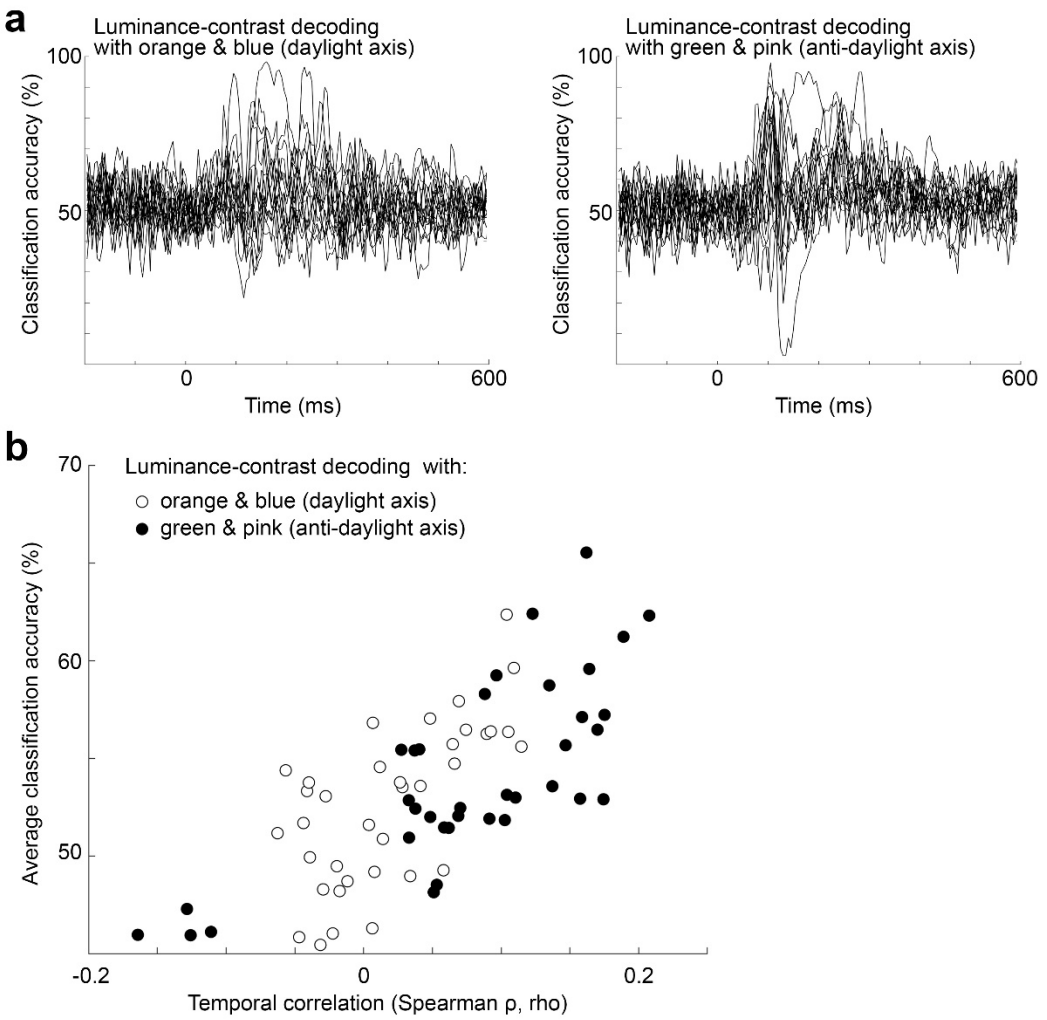
SI Figure 3. Data reliability analysis for main experiments. **a)** Results for the identity problems. For each identity problem, pairs of classifiers were trained and tested on independent samples containing 10%, 25%, 40%, and 50% of the data, and the correlation between the classifiers' performance 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. **b)** Results for the generalization problems. The test-retest curves in panel (a) differ somewhat from those in SI Figure 2, which shows the same analysis for the pilot experiment. The data for the main experiments were obtained in 18 participants using 8 colors, whereas the data for the pilot experiment were obtained in four participants using 4 colors.



SI Figure 4. The average classification accuracy for the 6 sets of hue identity problems between stimuli of the same luminance contrast (six problems among light hues; six problems among dark hues). The time to peak was identical for the two sets of problems (vertical line, 120 ms [120, 125]), and the amplitude of peak decoding was not significantly different (Light Hues: 0.7132 [0.6617, 0.7618]; Dark Hues: 0.7547 [0.7009, 0.8050]). The horizontal line shows time bins at which decoding was significantly above chance, FDR corrected, for a minimum of 5 consecutive time bins. We examined the time course for these two sets of problems to assess the impact of the potential difference in contrast between dark and light stimuli (the light and dark stimuli differ in contrast when measured in units of detection threshold; see SI Figure 1). These results show that this potential difference in contrast did not impact the time course of hue decoding.



SI Figure 5. Behavioral performance in the 1-back hue-matching task, and the 1-back luminance matching task for the two subjects who participated in the control experiment in which MEG data were collected while participants performed either the 1-back luminance-contrast matching task (black traces) or the 1-back hue matching task (blue traces). Behavioral responses were obtained during the sessions in which MEG data were collected (see **Figure 4** for the MEG decoding results of this control experiment). Note the slightly longer reaction times for the luminance task (Subject 1) and the slightly lower accuracy for the luminance task (Subject 2), which reflects the general difficulty people face in making heterochromatic luminance-contrast matches.



SI Figure 6. Luminance-contrast decoding was more variable across participants for problems using orange and blue versus problems using green and pink. **a)** Individual luminance-contrast decoding curves for problems involving orange and blue (left panel) and pink and green (right panel) across the 18 participants (36 problems illustrated in each panel, two per participant involving orange and blue; and two per participant involving pink and green). The decoding curves using pink and green were more correlated across time, as quantified in panel (b). **b)** The average classification accuracy (averaged over the entire decoding time course from 0 to 600 ms) for each problem, versus the temporal correlation of the problem. The temporal correlation for each problem was computed by measuring the correlation between it and a problem drawn at random from the remaining 71 problems, repeated 1000x, and averaged. Decoding luminance contrast was more variable for data obtained using orange and blue compared to data obtained with green and pink (the black dots are shifted to the right; MANOVA, p=0.0007).

References

1. Kuehni, R. & Schwartz, A. Color Ordered: A Survey of Color Systems from Antiquity to the Present. *New York, NY: Oxford University press* (2008).
2. Smet, K.A., Webster, M.A. & Whitehead, L.A. A simple principled approach for modeling and understanding uniform color metrics. *J Opt Soc Am A Opt Image Sci Vis* **33**, A319-331 (2016).
3. Ennis, R.J. & Zaidi, Q. Geometrical structure of perceptual color space: Mental representations and adaptation invariance. *J Vis* **19**, 1 (2019).
4. Favreau, O.E. & Cavanagh, P. Color and luminance: independent frequency shifts. *Science* **212**, 831-832 (1981).
5. Krauskopf, J., Williams, D.R. & Heeley, D.W. Cardinal directions of color space. *Vision Research* **22**, 1123-1131 (1982).
6. Goddard, E., Chang, D.H.F., Hess, R.F. & Mullen, K.T. Color contrast adaptation: fMRI fails to predict behavioral adaptation. *Neuroimage* **201**, 116032 (2019).
7. Krizhevsky, A., Sutskever, I. & Hinton, G.E. Imagenet classification with deep convolutional neural networks. *in Advances in Neural Information Processing systems*, 1097-1105 (2012).
8. Zeiler, M.D. & Fergus, R. Visualizing and Understanding Convolutional Networks. *Computer Vision - Eccv 2014, Pt I* **8689**, 818-833 (2014).
9. Rafegas, I. & Vanrell, M. Color representation in CNNs: parallelisms with biological vision. *Ieee Int Conf Comp V*, 2697-2705 (2017).
10. Flachot, A. & Gegenfurtner, K.R. Processing of chromatic information in a deep convolutional neural network. *J Opt Soc Am A Opt Image Sci Vis* **35**, B334-B346 (2018).
11. Hansen, T. & Gegenfurtner, K.R. Independence of color and luminance edges in natural scenes. *Vis Neurosci* **26**, 35-49 (2009).
12. Sachtler, W.L. & Zaidi, Q. Chromatic and luminance signals in visual memory. *J Opt Soc Am A* **9**, 877-894 (1992).
13. Clifford, C.W., Spehar, B., Solomon, S.G., Martin, P.R. & Zaidi, Q. Interactions between color and luminance in the perception of orientation. *J Vis* **3**, 106-115 (2003).
14. Yaguchi, H. & Ikeda, M. Subadditivity and superadditivity in heterochromatic brightness matching. *Vision Res* **23**, 1711-1718 (1983).
15. Buck, S.L. Brown. *Curr Biol* **25**, R536-537 (2015).
16. Lindsey, D.T. & Brown, A.M. Universality of color names. *Proc Natl Acad Sci U S A* **103**, 16608-16613 (2006).
17. De Valois, K.K. & Switkes, E. Simultaneous masking interactions between chromatic and luminance gratings. *Journal of the Optical Society of America* **73**, 11-18 (1983).
18. Wiesel, T.N. & Hubel, D.H. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J. Neurophysiol.* **29**, 1115-1156 (1966).
19. Reid, R.C. & Shapley, R.M. Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus. *J Neurosci* **22**, 6158-6175 (2002).
20. van der Horst, G.J. & Bouman, M.A. Spatiotemporal chromaticity discrimination. *J Opt Soc Am* **59**, 1482-1488 (1969).
21. Granger, E.M. & Heurtley, J.C. Letters to the editor: Visual chromaticity-modulation transfer function. *J Opt Soc Am* **63**, 1173-1174 (1973).
22. Mullen, K.T. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *Journal of Physiology* **359**, 381-400 (1985).
23. Lee, B.B., Martin, P.R. & Valberg, A. Sensitivity of macaque retinal ganglion cells to chromatic and luminance flicker. *Journal of Physiology* **414**, 223-243 (1989).
24. Martin, P.R., White, A.J., Goodchild, A.K., Wilder, H.D. & Sefton, A.E. Evidence that blue-on cells are part of the third geniculocortical pathway in primates. *Eur J Neurosci* **9**, 1536-1541 (1997).
25. Dobkins, K. Moving colors in the lime light. *Neuron* **25**, 15-18 (2000).
26. Maunsell, J.H*., et al.* Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Vis Neurosci* **16**, 1-14 (1999).
27. Gegenfurtner, K.R. Cortical mechanisms of colour vision. *Nat Rev Neurosci* **4**, 563-572 (2003).
28. Thorell, L.G., De Valois, R.L. & Albrecht, D.G. Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research* **24**, 751-769 (1984).
29. Nealey, T.A. & Maunsell, J.H. Magnocellular and parvocellular contributions to the responses of neurons in macaque striate cortex. *J Neurosci* **14**, 2069-2079 (1994).
30. Johnson, E.N., Hawken, M.J. & Shapley, R. Cone inputs in macaque primary visual cortex. *J Neurophysiol* **91**, 2501-2514 (2004).
31. Garg, A.K., Li, P., Rashid, M.S. & Callaway, E.M. Color and orientation are jointly coded and spatially organized in primate primary visual cortex. *Science* **364**, 1275-1279 (2019).

32. Lennie, P., Krauskopf, J. & Sclar, G. Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience* **10**, 649-669 (1990).
33. Conway, B.R. Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *Journal of Neuroscience* **21**, 2768-2783 (2001).
34. Horwitz, G.D. & Hass, C.A. Nonlinear analysis of macaque V1 color tuning reveals cardinal directions for cortical color processing. *Nat Neurosci* **15**, 913-919 (2012).
35. Nassi, J.J. & Callaway, E.M. Parallel processing strategies of the primate visual system. *Nat Rev Neurosci* **10**, 360-372 (2009).
36. Rabin, J., Switkes, E., Crognale, M., Schneck, M.E. & Adams, A.J. Visual evoked potentials in three-dimensional color space: correlates of spatio-chromatic processing. *Vision Res* **34**, 2657-2671 (1994).
37. Kuriki, I., Sadamoto, K. & Takeda, T. MEG recording from the human ventro-occipital cortex in response to isoluminant color stimulation. *Vis Neurosci* **22**, 283-293 (2005).
38. Logothetis, N.K., Schiller, P.H., Charles, E.R. & Hurlbert, A.C. Perceptual deficits and the activity of the color-opponent and broad-band pathways at isoluminance. *Science* **247**, 214-217 (1990).
39. Shevell, S.K. & Kingdom, F.A. Color in complex scenes. *Annu Rev Psychol* **59**, 143-166 (2008).
40. Engel, S.A. Adaptation of Oriented and Unoriented Color-Selective Neurons in Human Visual Areas. *Neuron* **45**, 613-623 (2005).
41. Wade, A., Augath, M., Logothetis, N. & Wandell, B. fMRI measurements of color in macaque and human. *J Vis* **8**, 6 1-19 (2008).
42. Mullen, K.T., Thompson, B. & Hess, R.F. Responses of the human visual cortex and LGN to achromatic and chromatic temporal modulations: an fMRI study. *J Vis* **10**, 13 (2010).
43. Boynton, G.M., Engel, S.A. & Heeger, D.J. Linear systems analysis of the fMRI signal. *Neuroimage* **62**, 975-984 (2012).
44. Carlson, T., Tovar, D.A., Alink, A. & Kriegeskorte, N. Representational dynamics of object vision: the first 1000 ms. *J Vis* **13** (2013).
45. van de Nieuwenhuijzen, M.E.*, et al.* MEG-based decoding of the spatiotemporal dynamics of visual category perception. *Neuroimage* **83**, 1063-1073 (2013).
46. Cichy, R.M., Pantazis, D. & Oliva, A. Resolving human object recognition in space and time. *Nat Neurosci* **17**, 455-462 (2014).
47. Isik, L., Meyers, E.M., Leibo, J.Z. & Poggio, T. The dynamics of invariant object recognition in the human visual system. *J Neurophysiol* **111**, 91-102 (2014).
48. Wardle, S.G., Kriegeskorte, N., Grootswagers, T., Khaligh-Razavi, S.M. & Carlson, T.A. Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. *Neuroimage* **132**, 59-70 (2016).
49. Hermann, K., Pantazis, D. & Conway, B.R. The dynamics of color processing in humans measured with MEG. *Society for Neuroscience Annual Meeting. 790.03/M32* (2015).
50. Rosenthal, I.A., Hermann, K.L., Vonder Haar, C., Pantazis, D. & Conway, B.R. Decoding hue and luminance with magnetoencephalography Society for Neuroscience annual meeting, 774.03. (2017).
51. Teichmann, L., Grootswagers, T., Carlson, T.A. & Rich, A.N. Seeing versus knowing: The temporal dynamics of real and implied colour processing in the human brain. *Neuroimage* **200**, 373-381 (2019).
52. Sandhaeger, F., von Nicolai, C., Miller, E.K. & Siegel, M. Monkey EEG links neuronal color and motion information across species and scales. *Elife* **8** (2019).
53. Hermann, K., Rosenthal, I., Singh, S., Pantazis, D. & Conway, B.R. Temporal dynamics of the neural mechanisms for encoding hue and luminance contrast uncovered by magnetoencephalography. *BioRxiv* **https://doi.org/10.1101/2020.06.17.155713** (2020).
54. Rosenthal, I., Singh, S., Hermann, K., Pantazis, D. & Conway, B.R. Uncovering the geometry of color space with magnetoencephalography (MEG). *BioRxiv* **https://doi.org/10.1101/2020.08.10.245324** (2020).
55. Hajonides, J.E., Nobre, A.C., van Ede, F. & Stokes, M.G. Decoding visual colour from scalp electroencephalography measurements. *BioRxiv* **https://doi.org/10.1101/2020.07.30.228437** (2020).
56. Rosenthal, I., Singh, S., Hermann, K., Pantazis, D. & Conway, B.R. Color space geometry uncovered with magnetoencephalography. *Curr Biol* **in press** (2020).
57. MacLeod, D.I. & Boynton, R.M. Chromaticity diagram showing cone excitation by stimuli of equal luminance. *J Opt Soc Am* **69**, 1183-1186 (1979).
58. Derrington, A.M., Krauskopf, J. & Lennie, P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology* **357**, 241-265 (1984).
59. Meyers, E.M. The neural decoding toolbox. *Front Neuroinform* **7**, 8 (2013).
60. Mannion, D.J., McDonald, J.S. & Clifford, C.W. Discrimination of the local orientation structure of spiral Glass patterns early in human visual cortex. *Neuroimage* **46**, 511-515 (2009).

61. Seymour, K., Clifford, C.W., Logothetis, N.K. & Bartels, A. Coding and binding of color and form in visual cortex. *Cereb Cortex* **20**, 1946-1954 (2010).
62. Brouwer, G.J. & Heeger, D.J. Categorical clustering of the neural representation of color. *J Neurosci* **33**, 15454-15465 (2013).
63. Norman-Haignere, S.V., Kanwisher, N., McDermott, J.H. & Conway, B.R. Divergence in the functional organization of human and macaque auditory cortex revealed by fMRI responses to harmonic tones. *Nature Neuroscience* **22**, 1057-+ (2019).
64. Switkes, E. & Crognale, M.A. Comparison of color and luminance contrast: apples versus oranges? *Vision Res* **39**, 1823-1831 (1999).
65. Hebart, M.N. & Baker, C.I. Deconstructing multivariate decoding for the study of brain function. *Neuroimage* **180**, 4-18 (2018).
66. Marti, S. & Dehaene, S. Discrete and continuous mechanisms of temporal selection in rapid visual streams. *Nat Commun* **8**, 1955 (2017).
67. Quentin, R.*, et al.* Differential Brain Mechanisms of Selection and Maintenance of Information during Working Memory. *J Neurosci* **39**, 3728-3740 (2019).
68. Hebart, M.N., Bankson, B.B., Harel, A., Baker, C.I. & Cichy, R.M. The representational dynamics of task and object processing in humans. *Elife* **7** (2018).
69. King, J.R. & Dehaene, S. Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci* **18**, 203-210 (2014).
70. Chauhan, T.*, et al.* The achromatic locus: effect of navigation direction in color space. *J Vis* **14** (2014).
71. Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G.D. & Hurlbert, A. Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *PLoS One* **9**, e87989 (2014).
72. Lafer-Sousa, R., Hermann, K.L. & Conway, B.R. Striking individual differences in color perception uncovered by 'the dress' photograph. *Curr Biol* **25**, R545-546 (2015).
73. Winkler, A.D., Spillmann, L., Werner, J.S. & Webster, M.A. Asymmetries in blue-yellow color perception and in the color of 'the dress'. *Curr Biol* **25**, R547-548 (2015).
74. Gegenfurtner, K.R., Bloj, M. & Toscani, M. The many colours of 'the dress'. *Curr Biol* **25**, R543-R544 (2015).
75. Brainard, D.H. & Hurlbert, A.C. Colour Vision: Understanding #TheDress. *Curr Biol* **25**, R551-554 (2015).
76. Goddard, E. A step toward understanding the human ventral visual pathway. *J Neurophysiol* **117**, 872-875 (2017).
77. Lafer-Sousa, R., Conway, B.R. & Kanwisher, N.G. Color-Biased Regions of the Ventral Visual Pathway Lie between Face- and Place-Selective Regions in Humans, as in Macaques. *J Neurosci* **36**, 1682-1697 (2016).
78. Desikan, R.S.*, et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968-980 (2006).
79. Toga, A.W., Thompson, P.M., Mori, S., Amunts, K. & Zilles, K. Towards multimodal atlases of the human brain. *Nat Rev Neurosci* **7**, 952-966 (2006).
80. Haynes, J.D. & Rees, G. Decoding mental states from brain activity in humans. *Nat Rev Neurosci* **7**, 523-534 (2006).
81. Tong, F. & Pratte, M.S. Decoding patterns of human brain activity. *Annu Rev Psychol* **63**, 483-509 (2012).
82. Proklova, D., Kaiser, D. & Peelen, M.V. MEG sensor patterns reflect perceptual but not categorical similarity of animate and inanimate objects. *Neuroimage* **193**, 167-177 (2019).
83. Kaiser, D., Azzalini, D.C. & Peelen, M.V. Shape-independent object category responses revealed by MEG and fMRI decoding. *J Neurophysiol* **115**, 2246-2250 (2016).
84. Dobs, K., Isik, L., Pantazis, D. & Kanwisher, N. How face perception unfolds over time. *Nat Commun* **10**, 1258 (2019).
85. Lafer-Sousa, R. & Conway, B.R. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat Neurosci* **16**, 1870-1878 (2013).
86. Bannert, M.M. & Bartels, A. Human V4 Activity Patterns Predict Behavioral Performance in Imagery of Object Color. *Journal of Neuroscience* **38**, 3657-3668 (2018).
87. Siuda-Krzywicka, K., Witzel, C., Bartolomeo, P. & Cohen, L. Color Naming and Categorization Depend on Distinct Functional Brain Networks. *Cereb Cortex* (2020).
88. Conway, B.R. Color vision, cones, and color-coding in the cortex. *Neuroscientist* **15**, 274-290 (2009).
89. Roe, A.W.*, et al.* Towards a unified theory of visual area V4. *Neuron* **74**, 12-29 (2012).
90. Bohon, K.S., Hermann, K.L., Hansen, T. & Conway, B.R. Representation of Perceptual Color Space in Macaque Posterior Inferior Temporal Cortex (the V4 Complex). *eNeuro* **3** (2016).
91. Cauchoix, M., Barragan-Jason, G., Serre, T. & Barbeau, E.J. The neural dynamics of face detection in the wild revealed by MVPA. *J Neurosci* **34**, 846-854 (2014).
92. Goddard, E., Carlson, T.A., Dermody, N. & Woolgar, A. Representational dynamics of object recognition: Feedforward and feedback information flows. *Neuroimage* **128**, 385-397 (2016).

93. 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* **153**, 346-358 (2017).
94. Rathbun, D.L., Warland, D.K. & Usrey, W.M. Spike timing and information transmission at retinogeniculate synapses. *J Neurosci* **30**, 13558-13566 (2010).
95. Jin, X. & Costa, R.M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457-462 (2010).
96. King, J.-R. & Wyart, V. The Human Brain encodes a Chronicle of Visual Events at each Instant of Time. *bioRxiv*, 846576 (2019).
97. Reinagel, P. & Reid, R.C. Precise firing events are conserved across neurons. *J Neurosci* **22**, 6837-6841 (2002).
98. Hubel, D.H. & Wiesel, T.N. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* **195**, 215-243 (1968).
99. Johnson, E.N., Hawken, M.J. & Shapley, R. The orientation selectivity of color-responsive neurons in macaque V1. *J Neurosci* **28**, 8096-8106 (2008).
100. Lafer-Sousa, R. & Conway, B.R. #TheDress: Categorical perception of an ambiguous color image. *J Vis* **17**, 25 (2017).
101. Conway, B.R. The Organization and Operation of Inferior Temporal Cortex. *Annu Rev Vis Sci* **4**, 381-402 (2018).
102. Cicmil, N., Bridge, H., Parker, A.J., Woolrich, M.W. & Krug, K. Localization of MEG human brain responses to retinotopic visual stimuli with contrasting source reconstruction approaches. *Front Neurosci* **8**, 127 (2014).
103. 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).
104. Brainard, D.H. Cone contrast and opponent modulation color spaces in *Human Color Vision* (P.K. Kaiser and R.M. Boynton Editors, 2nd Edition). *Washington, D.C.: Optical Society of America* (1996).
105. Smith, V.C. & Pokorny, J. Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research* **15**, 161-171 (1975).
106. Kleiner, M., Brainard, D.H. & Pelli, D. What's new in Psychtoolbox-3? *Perception* **36** (2007).
107. Smith, V.C. & Pokorny, J. Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research* **15**, 161-171 (1975).