Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Decoding of EEG signals reveals non-uniformities in the neural geometry of colour

Tushar Chauhan^{a,*}, Ivana Jakovljev^b, Lindsay N. Thompson^c, Sophie M. Wuerger^d, Jasna Martinovic^{c,e,*}

^a The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 02139 Cambridge MA, USA

^b Department of Psychology. Faculty of Philosophy, University of Novi Sad, Serbia

^c School of Psychology, University of Aberdeen, Aberdeen, AB24 3FX, UK

^d Department of Psychology, University of Liverpool, Liverpool, L697ZA, UK

e Department of Psychology, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, EH8 9JZ, UK

ARTICLE INFO

Keywords: Unique hues Electroencephalography Decoding Population encoding Colour perception

ABSTRACT

The idea of colour opponency maintains that colour vision arises through the comparison of two chromatic mechanisms, red versus green and yellow versus blue. The four unique hues, red, green, blue, and yellow, are assumed to appear at the null points of these the two chromatic systems. Here we hypothesise that, if unique hues represent a tractable cortical state, they should elicit more robust activity compared to other, non-unique hues. We use a spatiotemporal decoding approach to report that electroencephalographic (EEG) responses carry robust information about the tested isoluminant unique hues within a 100–350 ms window from stimulus onset. Decoding is possible in both passive and active viewing tasks, but is compromised when concurrent high luminance contrast is added to the colour signals. For large hue-differences, the efficiency of hue decoding can be predicted by mutual distance in a nominally uniform perceptual colour space. However, for small perceptual neighbourhoods around unique hues, the encoding space shows pivotal non-uniformities which suggest that anisotropies in neurometric hue-spaces may reflect perceptual unique hues.

1. Introduction

The idea of colour opponency maintains that colour vision arises through the comparison of two chromatic mechanisms, red versus green (RG) and blue versus yellow (BY). The four unique hues, red, green, blue, and yellow, are assumed to appear at the null points of these the two chromatic systems (De Valois and De Valois, 1993; Hering, 1920; Jameson and Hurvich, 1964). Colour vision starts in the retina, where light is absorbed in receptors (long-, medium, and short-wavelength sensitive cone receptors - L, M, S) and small bistratified ganglion cells that receive S(M + L) cone input have been postulated to be the retinal origin of the BY channel, while midget ganglion cells that take the differences between the L and M cone output were believed to be the retinal origin of the RG channel (Lee et al., 2010). However, it has now been confirmed that the chromatic tuning of behaviourally characterised opponent channels differs from these early cone-opponent mechanisms, hence another transformation of chromatic signals must take place between the Lateral Geniculate Nucleus (LGN) and the primary or extrastriate visual cortex (De Valois and De Valois, 1993; Wuerger et al., 2005).

While some neuroimaging studies have attempted to identify a neural basis for unique hues, their results remain controversial. Stoughton and Conway (2008) reported neuronal clusters which were preferentially tuned to unique hues in the posterior inferior temporal (PIT) cortex of macaques. However, their findings have been challenged on the grounds that the study was not fully controlled for low-level differences in neuronal tuning, which could provide a more parsimonious explanation for their results (Bohon et al., 2016; Conway and Stoughton, 2009; Mollon, 2009). Similarly Forder et al. (2017a) reported that event-related potentials (ERPs) for unique hues show decreased latencies compared to non-unique hues. But the reported difference in peak latencies could, once again, have stemmed from differential activation of low-level, cone-opponent processes, to which ERPs are particularly sensitive (Knoblauch et al., 1998; Rabin et al., 1994). Thus, the neural basis of these cortical hue-opponent chromatic systems, and consequently, the unique hues, still remains an open problem.

* Abbreviated title: Decoding unique hues from EEG signals

* Corresponding authors.

E-mail addresses: tchauhan@mit.edu (T. Chauhan), jmartino@ed.ac.uk (J. Martinovic).

https://doi.org/10.1016/j.neuroimage.2023.119884.

Received 24 April 2022; Received in revised form 4 November 2022; Accepted 15 January 2023 Available online 16 January 2023.

1053-8119/© 2023 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)







One of the major reasons for the failure to address this issue has been the fact that neural activity is rich in coding possibilities which complicate our understanding of the relationship between external stimuli and the evoked response (Jazaveri and Afraz, 2017; Johnson, 2000). This is particularly true if potential low-level confounds can lead to a stronger, overlapping signal. This seems to be the case for unique hues, whose encoding is bound to overlap with, and be influenced by, the encoding of luminance contrast (for human fMRI see Goddard and Mullen, 2020; for macaque neurophysiology see Namima et al., 2014; for human EEG see Nunez et al., 2017). Ritchie et al. (2019) suggest that an ideal way to utilise neural decoding is to reconstruct an activation space from multivariate neural data and make psychological inferences by assessing whether such activation spaces correspond to psychological constructs. Recent studies have begun to apply this approach to challenges in colour neuroscience such as identifying the neural representations that underlie colour geometries (Rosenthal et al., 2021).

We hypothesise that, if there is indeed a distinct and discernible neural signature for unique hues, it should be reflected in the structure of the neurometric hue-representational space described by EEG signals. We used a decoding paradigm to test this hypothesis in two stages. First, we demonstrate that under isoluminant conditions, hue information can indeed be extracted from EEG signals, and that crucially, the encoding for unique hues is more robust than non-unique hues. To establish that our predictions generalise beyond a single decoding context (stimulus or task-wise), we test our decoding prediction using both active and passive viewing tasks. Second, we show that the structure of the neurometric space which encodes hue is distorted in the local neighbourhood of unique hue representations - suggesting an anisotropic mapping between perceptual colour and its cortical representation. Taken together, our findings suggest that the neural basis of perceptual unique hues may reside in a set of stable fixed-points of a spatiotemporal population code for colour representations in the cortex.

2. Methods

2.1. Participants

In Experiment 1, twenty participants (16 females, 4 males) completed the study, ranging in age from 18 to 38 y.o. a. (mean age 21 y.o.a). In Experiment 2, 16 participants (all female) completed the study, ranging in age 19 – 32 years old (mean age 22 years). All participants reported normal or corrected-to-normal visual acuity, In Experiment 1, their colour vision was verified using the Trivector Cambridge Colour Test (Regan et al., 1994). In Experiment 2, we relied on the City University Colour Vision Test (Fletcher, 1975). Participants gave written informed consent and were reimbursed for their effort and time. The study was approved by the ethics committee of the School of Psychology, University of Aberdeen, and was in accordance with the Declaration of Helsinki.

2.2. Stimuli

The experiments were programmed using the CRS Toolbox and color Toolbox (CRS, UK) for MATLAB (Mathworks, USA). In Experiment 1, stimuli were rendered on a 21-inch Viewsonic P227F CRT Monitor which was placed 70 cm away from the participant. The monitor was controlled through a Visage system (CRS, UK) and calibrated using ColorCAL2 (CRS, UK). Colours were generated on the basis of measurements taken with a SpectroCAL (CRS, UK). Participants gave their responses using a Cedrus R530 response box (San Pedro, USA). In Experiment 2, colours were presented on a Display++ (CRS, UK) device, and responses were recorded using a CT-6 button box (CRS, UK).

Different sets of colours were used in the two experiments. In Experiment 1, stimulus colours were selected from a large, normative dataset of unique hues (Wuerger and Xiao, 2015). Fig. 1B shows the coordinates of the stimuli in CIE 1976 Uniform Colour Space (CIE 1976 UCS; see Schanda, 2016), while Supplementary Table S1 lists the coordinates in cone-activation space. The hue angles for unique red (UR) and unique green (UG) stimuli corresponded to mean values in the dataset - in CIE 1976 UCS, the angles were 14.4° and 133.4° for UR and UG respectively. Orange and turquoise stimuli were chosen such that they bisected the hue angles between the two adjacent unique hues. Orange (hue angle 41.5°) was the intermediate hue between UR and unique yellow, and turquoise (hue angle 185.1°) was the intermediate hue between UG and unique blue. While there is indeed variation in individual unique hue settings, the choice of stimuli in this experiment was motivated by the fact that intra-observer variability in unique hues is consistently reported to be much lower than inter-observer variability. For instance, Xiao et al. (2011) report intra-observer variability to be about half the inter-observer variability, while Hinks et al. (2007) report this ratio to be about 0.15. Meanwhile, the intermediate hues were chosen to fall as far away as possible from normative unique hue values (i.e., intermediate to them). Thus, the chosen stimulus colours were suitable for investigating asymmetries in the neural processing of hue for large colour distances. The perceptual uniformity of the CIE 1976 UCS colour space allows us to quantify the distances between pairs of stimulus colours. Here, they are: 27.1° between red and orange, 51.7° between green and turquoise, 91.9° between orange and green, 119° between red and green, 170.7° between red and turquoise and 143.6° between orange and turquoise. Thus, average distance from the two neighbouring colours is: 98.9° for red, 59.5° for orange, 71.8° for green and 111.2° for turquoise. Three stimulus luminance levels were used: nominal isoluminance (24 cd/m²), 45% Weber contrast (34.8 cd/m²), and 90% Weber contrast (45.6 cd/m²). We used the same CIE 1976 UCS coordinates for a given colour at all luminance values, ensuring that the colours at each luminance level were equally saturated in the CIE 1976 UCS plane. The background had the following CIE 1931 xyY coordinates (Smith and Guild, 1931): 0.3127, $0.3290, 24 \ cd/m^2.$

In Experiement 2, stimuli were based on each participant's individual settings for two unique hues (yellow and green) and two intermediate hues (orange and turquoise). For a given observer, the final set of 12 stimulus colours consisted of the average individual setting for each hue as well as hues situated 10° to the left and right of these individual settings in CIELCh colour space (Fig. 1C; see Supplementary Figure S1 for the coordinates in cone-activation space). In the EEG experiment, these hues were presented at a Lightness value of 45 and a Chroma value of 25. The just noticeable difference (JND) in this region of the CIELAB space is ~ 2.3 ΔE_{LAB} (CIE, 2004; Fairchild, 2013). At this level of chroma, a difference of 10° in hue is equivalent to ~ 4.36 ΔE_{LAB} , i.e., ~ 1.89 JNDs. Thus, within each triplet of neighbouring colours, the hues were discriminable but remained highly similar to each other. With these colour values, we effectively had four clusters of colours corresponding to the hues orange, yellow, green, and turquoise for each participant, with each cluster consisting of the individual setting for that hue, along with two flanking colours $\pm 10^{\circ}$ from the setting (e.g., unique yellow, a yellow 10° counter-clockwise and a yellow 10° clockwise). All colours were nominally isoluminant with the background (CIE 1931 xyY coordinates: 0.3127, 0.3290, 22.93 cd/m²).

2.3. Procedure

EEG data was recorded during a shape discrimination task. The purpose of the task was to engage participants' attention in a stimulus dimension orthogonal to colour - i.e., shape. The stimuli consisted of uniformly coloured shapes shown against a grey background. Each trial began with the appearance of a fixation cross, followed by a 2-degree circular stimulus (passive viewing event) which changed shape (shape change event) into either a diamond or a square (Fig. 1). The passive viewing event occurred 700 \pm 200 ms after the appearance of the fixation cross, and the shape change event occurred 800–1500 ms after the passive viewing event.



Fig. 1. Experimental design. A. Trial design. Each trial started with the appearance of a fixation cross, which was followed by the presentation of a circular uniformly-coloured stimulus at a random offset of 700±200 ms. At a random time-point 800-1500 ms after stimulus onset, the shape of the stimulus changed from circular to either square or diamond. Participants were instructed to discriminate the final shape via a button press as quickly and as accurately as they could. Each trial ended 2 s after stimulus onset. Two events were defined during each trial: a passive viewing event defined by the appearance of the stimulus, and a shape change event defined by the change in stimulus shape. B. Stimulus hues for Experiment 1. The stimuli were one of four hues: red (R), orange (O), green (G) or Turquoise (T), shown as coloured discs of the corresponding hue. The abscissa (u') and the ordinate (v')in the plot denote coordinates in the CIE 1976 uniform chromaticity space. The orientations of mean unique hues from Xiao et al. (2011) are also shown using translucent lines of the corresponding colour (unique red: red line, unique yellow: yellow line, etc.) passing through the grey background (grey disc). The R and G stimuli were unique hues, while the O and T stimuli were chosen such that their hue angle bisected the nearest unique hues (O bisects unique red and unique yellow; T bisects unique green and unique blue). C. Stimuli for Experiment 2. In a psychophysical experiment before the EEG recordings, the observer settings for four hues (orange: O, yellow: Y, green: G, and turquoise: T) were measured using the method of adjustment (see Methods). Each observer was then presented with the mean of their respective settings as well as hues 10° clockwise and anti-clockwise from

u' their settings. Each observer's settings (=) are snown as coloured discs of the corresponding hue (orange disc for O, yellow disc for Y, etc.). The clockwise rotations (-) are shown as darker discs of the corresponding hue, while the anticlockwise rotations (+) are shown as lighter discs. The average settings for each stimulus condition are shown as large diamond symbols of the corresponding colour. The background is shown as a grey disc. Like B, the CIE 1976 uniform chromaticity space is used.

Participants identified the final shape of the stimulus using the left or the right button on a button box. The assignment of button to the target shape was counterbalanced between participants. The conditions were randomly intermixed, with a different order for each participant. The entire experiment was conducted in a sound-attenuated, electrically shielded chamber, with the screen being the only source of light. In addition to EEG recordings (described below), two other task-related variables were measured – task accuracy and reaction time. For each colour and shape combination, we had 30 trials. As diamond and square shape-change trials were subsequently collapsed together, this resulted in 60 trials per colour and 720 trials in total, presented in random order and divided into 10 blocks. This was the same for both experiments. In addition, Experiment 1 was preceded by a practice of 24 trials, while Experiment 2 was preceded by a practice of 16 trials. The EEG task took approximately 50 min to complete.

After the completion of the EEG experiment, participants rated each colour on a 9-point Likert scale for the representativeness of its category. Participants were asked to imagine the perfect representative for a colour category and rate how representative a sample was of that category, with 1 being the least representative and 9 being the most representative. All colours were displayed simultaneously on the screen during this procedure and remained on the screen until the participants completed the task. Colours were presented on the computer screen as a set of 4 rows of squares that showed the three luminance (Experiment 1) or hue (Experiment 2) values for that colour. Participants completed the task in approximately 5 min. Note that these ratings were included to provide data on the proximity of each stimulus colour to its focal colour, i.e., the best example of its category and thus did not relate to its unique hue status. We used the ratings to conduct a control analysis

to understand if the neural representations identified through information decoding related in any way to colour categories (for previous work on colour categories and EEG, see Clifford et al., 2010; Fonteneau and Davidoff, 2007; Forder et al., 2017b; Holmes et al., 2009; Thierry et al., 2009).

There were also two additional measures, specific to each experiment. In Experiment 1, for each participant, heterochromatic flicker photometry (HCFP) at 20 Hz (Walsh, 1958) was used to establish the departure from isoluminance for all colours. The task required the participant to adjust the luminance of the colour until perceived flicker was minimised. Participants performed 8 trials per colour – the step size was 0.5 cd/m2 and the flicker started from a randomly determined point that could be five steps above or below nominal isoluminance. These measurements were conducted to evaluate any individual differences in the amount of luminance contrast effectively present in nominally isoluminant stimuli. Rabin et al. (1994) demonstrate that departures from isoluminance need to be substantial to influence chromatic visual evoked potentials. Collecting HCFP data enabled us to verify that small departures from effective luminance did not significantly influence the efficiency of colour decoding.

Experiment 2 began with a hue adjustment task, in which participants made their individual hue settings for two unique hues (yellow and green) and two intermediate, non-unique hues (orange and turquoise). Participants performed one block of eight trials for each hue. The order of blocks (yellow, green, orange, turquoise) was randomized for each participant. Colours were defined in CIE LCh colour space to have the same chroma (C = 25) and lightness (L = 55). Initial hue angles were randomised to the following values: 90°+/- 12° for yellow, 180°+/- 12° for green, 45°+/- 12° for orange and 225°+/- 12° for turquoise. A coloured

2° circle was shown in the middle of the computer screen. Participants used the right and left buttons to change the hue along the CIE LCh hue circle in steps of 2° clockwise and counter-clockwise, respectively. Once the participants were happy with their setting, they completed the adjustment by pressing the top button. The task took approximately 10 min to complete. The first 6 participants performed the task without context. For the following 13 participants, we also presented a colour palette consisting of 19 squares 1° in size that ranged +/- 45° around the initial hue value, in steps of 5° of hue angle, positioned at 8.22° above the central stimulus. The colour palette provided context for the hue setting task. A between-subject ANOVA showed no difference in unique hue settings with and without context (F(1,14) = 0.23; p = .64, $\eta p^2 = 0.02$).

In total, the experiments lasted between two and a half and three hours, including the time to set up and remove the EEG electrodes.

2.4. EEG recording and pre-processing

Continuous brain activity was recorded from 64 scalp locations using active Ag-AgCl electrodes and 4 ocular channels (providing VEOG and HEOG) connected to a BioSemi Active-Two amplifier system (BioSemi, Amsterdam, The Netherlands) at a sampling rate of 256 Hz. Data processing was performed using EEGLAB (Delorme and Makeig, 2004) for Matlab (Mathworks, UK). Epochs lasting 900 ms were extracted: 200 ms before the relevant event (stimulus onset or shape change) and 700 ms afterwards. Data was low pass filtered at 40 Hz using a sinc FIR filter with a Kaiser window whose beta parameter was set to 5.653 (this is similar to a Henning window; Widmann et al., 2015). All trials with incorrect answers were excluded prior to the analysis. artefact removal was then performed by using the FASTER toolbox (Nolan et al., 2010), the ADJUST toolbox (Mognon et al., 2011), and self-written procedures in MATLAB. FASTER is an automated procedure that detects contaminated trials and noisy channels that need interpolation (either in the entire EEG recording or on any single trials) by calculating statistical parameters of the data and using a z-score of ± 3 as the metric that defined contaminated data. ADJUST is an automated procedure that operates on maps resulting from independent component analysis of EEG data, using properties of these components to label them as eye blinks, vertical or horizontal eye movements, or channel discontinuities so that they can be subtracted from the recording. We first rejected trials with global artifacts using FASTER, then ran an independent component analysis and applied ADJUST to the obtained decompositions, and finally, conducted channel interpolation with FASTER. In addition, any trials with eye movements were rejected based on $\pm 25 \mu V$ deviations from the horizontal electrooculogram in the uncorrected data. Blinks were rejected using a thresholding procedure similar to FASTER (Junghöfer et al., 2000).

Incorrect and rejected trials amounted to a very small proportion of the data – in Experiment 1, between 1% and 13% of total trials, and in Experiment 2, between 3% and 17% of total trials.

2.5. EEG classification

The classification of EEG signals was set up as a set of time-windowed error-correcting output codes models (tECOC) operating on 20 ms snippets of the signals (other reasonable time-windows yielded similar results, see Supplementary **Figure S2A**) from the occipital electrodes (the entire set of 64 electrodes yielded similar results, see Supplementary **Figure S2B**). Linear discriminant analysis (LDA) classifiers were employed as learning units due to their relative simplicity and computational efficiency. Denoting the EEG activity as a random multivariate variable X, and the stimulus label (colour and/or luminance) by the random variable Y (where realisations of Y are drawn from the set of all possible labels denoted L), the probability that the observed activity x is elicited by the stimulus described by label y is given by the Bayes

rule:

$$P(Y = y | X = x) = \frac{P(X = x | Y = y)P(Y = y)}{\sum_{l \in L} P(X = x | Y = l)P(Y = l)}$$

In LDA, the likelihood term is estimated by a multivariate Gaussian density function:

$$P(X = x|Y = y) = \frac{1}{\sqrt{(2\pi)^{N_e} |\Sigma|}} e^{-\frac{1}{2} (x - \mu_y)^T \Sigma^{-1} (x - \mu_y)}$$

Here, N_e is the number of electrodes, μ_y is the mean EEG activity for the label *y*, and Σ is the covariance matrix of the activity. The logposterior objective function $\delta_y(\mathbf{x})$ for the label *y* can thus be written as:

$$\delta_{y}(\boldsymbol{x}) = \log P(Y = y) - \frac{1}{2}\boldsymbol{\mu}_{y}^{T}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{y} + \boldsymbol{x}^{T}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{y}$$

Data for each observer was modelled separately, and the whole process was repeated 10 times. In each repetition for any given observer, the data were split into 5 folds containing roughly equal number of samples for each label. Each of the five folds was then tested by training the model on the remaining 4 folds. The entire pipeline was repeated 10 times for each observer. tECOC analysis gave us a time-series of confusion matrices (CMs) which characterise the model performance over the duration of the trial (see Supplementary Video V1). At each timepoint, while the diagonal of the CM gives a measure of model accuracy (true positive rate), the off-diagonal elements represent misclassifications, which are crucial towards understanding the topography and information content of the representational space (see Representational Similarity Analysis below). 95% confidence intervals, reported as shaded regions around the mean, were calculated using a two-tailed non-parametric permutation test (1000 samples were drawn). In addition to reporting the intrinsic variability in model performance, a comparison with a randomised model was also made (shown as horizontal lines under the relevant graphs). In each case, the randomised model was trained using a shuffled set of labels to estimate empirical chance performance, and the performance of the randomised model was compared with the actual model using a two-tailed randomisation test with 1000 permutations. Furthermore, when considered on its own, the performance of the randomised models was found to be close to theoretical chance level under the assumption of equilikelihood (see Supplementary Figure S2C).

2.6. Representational similarity analysis

The time-series of confusion matrices estimated by tECOC models were used to calculate pairwise dissimilarities between stimulus classes. Given a confusion matrix *C*, where each element c_{ij} denotes the probability of the stimulus type *i* being labelled as *j* by the model, first, a label-normalised matrix *S* was constructed such that $s_{ij} = c_{ij}/c_{ii}$. This asymmetric measure was then used to calculate a symmetric dissimilarity tensor Δ_{tECOC} given by

$$\Delta_{\text{tECOC}} = 1 - \max\left(0, 1 - \sqrt{SS^T}\right) \tag{1}$$

Here, the geometric mean across stimulus pairs is used to generalise distances in representational space (Kaneshiro et al., 2015; Shepard, 1958). A similar estimation was also made for perceptual data by considering pairwise absolute differences in CIELAB hue angles of the stimuli. These differences were used to estimate a perceptual dissimilarity matrix by first normalising across the rows to get a local distance measure which summed to one (similar to the normalisation over rows of the confusion matrix, e.g., in Kaneshiro et al., 2015), and then calculating the symmetrical dissimilarity matrix using Eq. (1). The perceptual dissimilarity was compared to Δ_{tECOC} using rank-correlation estimates (Kendall's coefficient).

2.7. Data and code availability

The decoding scripts have been packaged as the tECOC toolbox, which has been made available as a public git repository here. The EEG and behavioural data from both experiments have been shared on the Open Science Framework website here.

3. Results

3.1. Experiment 1: decoding unique and intermediate hues with and without luminance contrast

We measured EEG signals in a cohort of 20 participants while they viewed coloured stimuli (coloured shapes on a grey background) consisting of two unique hues – unique green and unique red, and two non-unique hues – orange and turquoise. In each trial, a coloured disc changed shape to a diamond or a square at a random time-point 800–1500 ms after stimulus onset (Fig. 1). The participant's task was to identify the target shape. The stimuli were either isoluminant with the background (0% luminance contrast), or presented at 45% or 90% luminance-contrast. This gave us a dataset of EEG signals labelled both in hue and luminance-contrast.

The task was easy, resulting in high overall accuracy ($95\% \pm 1\%$ SE, see Supplementary **Figure S3A**) and very fast responses (mean response time (RT) of 462 ± 15 ms, see Supplementary **Figure S3B**). Response-time data were analysed with a 3×4 repeated measures ANOVA (3 levels of luminance contrast vs. 4 hues), which yielded a significant main effect of luminance contrast (F(1.49, 28.28) = 67.56, p < 0.001, $\mu p^2 = 0.78$) and interaction with hue (F(6, 114) = 3.56, p = 0.003, $\mu p^2 = 0.16$) – hue itself did not have an effect (F(1.88, 35.79) = 2.93, p = 0.07). We deconstructed the interaction by performing separate repeated measures ANOVAs at each luminance contrast: while at isoluminance there was a significant effect (F(3, 57) = 6.19, p = 0.001, $\mu p^2 = 0.25$) driven by slower RTs for green (vs. red P = 0.019; vs orange P = 0.008, vs. turquoise P = 0.003), there were no differences at 45% luminance contrast (p = 0.16) or at 90% luminance contrast (p = 0.11).

After the completion of the EEG experiment, participants rated each colour on a 9-point Likert scale for its representativeness of its category (red, orange, green or turquoise). The average ratings and their SEs were as follows (see Supplementary **Figure S3C**): isoluminant red 4.35 ± 0.48 ; red at 45% luminance 2.85 ± 0.32 ; red at 90% luminance 1.90 ± 0.23 ; isoluminant green 7.70 ± 0.23 ; green at 45% luminance 6.10 ± 0.35 ; green at 90% luminance 5.55 ± 0.43 ; isoluminant orange 3.75 ± 0.48 ; orange at 45% luminance 4.15 ± 0.43 ; orange at 90% luminance 3.60 ± 0.32 ; isoluminant turquoise 6.00 ± 0.47 ; turquoise at 45% luminance 6.75 ± 0.38 ; turquoise at 90% luminance 6.40 ± 0.5 .

3.1.1. Unique hues can be robustly decoded from EEG signals

First, we asked whether the measured EEG waveforms contain consistent, discernible information about the hue of the stimulus. To do this, we trained tECOC models for each observer using only EEG responses to isoluminant stimuli, as this ensured minimal interference from luminance-contrast signals. In the first instance, we performed this analysis for epochs defined by the passive viewing event. We found that within a 100-300 ms window after stimulus onset, both unique hues could be decoded with above-chance accuracy (Fig. 2A). The nonunique hues, on the other hand, showed a much lower score (Fig. 2B). This pattern is stable over a range of tECOC time-windows (Supplementary Figure S2A) and also holds when the entire set of 64 electrodes is used (Supplementary Figure S2B). Furthermore, a bootstrapped power analysis shows high statistical power in the time-window of maximal discrimination (Supplementary Figure S4). The presence of signal on all electrodes is not surprising - unlike functional magnetic resonance imaging (fMRI), EEG does not detect localised physiological activity in a volume, but instead picks up a linear superposition of signals from a range of physiological sources. Thus, the signal is present in some degree at all sensors, with its amplitude (and thus also its signal to noise ratio) dependant on the position of the sensor relative to the source(s) (see, e.g., Maris, 2012 for a discussion of the so-called common pick-up problem).

For each participant, we also measured subjective isoluminance for each stimulus colour (see *Methods* for details). While one participant did not understand the task, the means, SEs and ranges of the settings from the remaining 19 participants were as follows: red 0.14 ± 0.57 cd/m^2 (-6 to $5.25 cd/m^2$); green $-1.09 \pm 0.49 cd/m^2$ (-6.58 to $1 cd/m^2$); orange 0.08 ± 0.56 (-4.34 to $6.50 cd/m^2$); turquoise (-0.05 $\pm 0.65 cd/m^2$).

Model accuracy quantifies the ability of the model to correctly identify the hue of a stimulus when presented with the corresponding EEG response. Theoretically, it is the sum of hit rates (true positive rates) for all labels, and corresponds to the diagonal of the confusion matrix. However, a deeper insight into model performance can be obtained when, in addition to the detection accuracy for a given input class, one also considers the probability of misclassification of inputs from this class. To investigate this, we estimated the off-diagonal elements of the confusion matrix. This allowed us to infer which classes are most likely to be confused by the model - thus providing a means of understanding how similar the information contained in EEG signals corresponding to different hues is. The subpanels of Fig. 2C (see also Supplementary Video V1) show the probability (over time) with which the model assigns each of the four hue labels to EEG responses elicited by a given input hue (the input hue is labelled above each subpanel). Thus, each subpanel in Fig. 2C shows one row of the confusion matrix. Within a 100-300 ms window, each input hue is only confused with its proximal pair (red and orange, and green and turquoise), while the prediction probabilities for non-proximal hues are below chance. This is also reflected in the checkerboard-like pattern observed in Supplementary Video V1. Furthermore, the model is likely to label EEG responses to non-unique hues (orange and turquoise) as being elicited by their proximal unique hues (red and green respectively) with almost equal probability, but not vice-versa. Once again, this suggests that EEG signals between 100 and 300 ms carry more robust representations of unique hues compared to non-unique hues.

The passive viewing at trial outset was followed by a change in the shape of the stimulus from a circle to either a square or a diamond at a random time-point 800-1500 ms from stimulus onset (see Fig. 1). The colour of the stimulus was task-irrelevant, and the hypothesis here was that since the observer will be attending to the stimulus shape, the EEG signal would be qualitatively different between the passive and shapechange segments. This would, in-turn, allow us to test if this difference is reflected in the ability of the model to classify hue information in the signal. It has been argued that colour-related activations should still be observed as long as the hue remains unattended and task-irrelevant (Forder et al., 2017b). To test this hypothesis, we trained tECOC models on the epochs defined by the shape-change event. As expected, the two segments were found to elicit activity which differed significantly both in the sequence of ERP peaks as well as topography (Fig. 3A). However, despite this difference, we were able to perform hue detection during the shape-change segment with an accuracy very similar to the passive viewing segment - both in terms of peak decoding score and its timecourse (Fig. 3B). This suggests that the temporal structure of the huerelated information in EEG signals is indeed robust to changes in the task (as long as the hue itself remains task-irrelevant), and can be extracted even when the observer is engaged in a concurrent shape discrimination task.

3.1.2. Luminance signals interfere with chromatic information in occipital ERPs

Next, we investigated whether hue identity could still be decoded when both chromatic and luminance information was present in the EEG signal. A chromatic-driven ERP is characterised by a robust negative



Fig. 2. Decoding isoluminant Unique and Non-unique hues from EEG responses. tECOC classification models were trained on EEG responses recorded in N = 20participants as they viewed isoluminant stimuli (Unique Hues: red and green; Non-unique hues: orange and turquoise). Note that all panels represent the results from a single set of models trained on isoluminant stimuli. A. Model accuracy for Unique hues. This corresponds to presenting the trained model with EEG responses to Unique Hue stimuli and estimating the probability with which the model is able to determine the correct stimulus hue (diagonal of the confusion matrix). The two solid lines show the mean accuracy of the model at each time-point. The hues are colour-coded, with the red and green lines representing model accuracy for unique red and unique green stimuli respectively. The shaded regions around the lines show bootstrapped 95% confidence intervals. A dashed line indicates the theoretical chance performance of the model (the empirical chance performance closely followed the theoretical chance level, and is shown in Supplementary Figure S2C). The two inlays show the classification accuracy (top-left: unique red, top-right: unique green) of models trained for each of the 20 observers. Only 100-300 ms after stimulus onset are shown in the inlays. The solid lines at the bottom show the period when the classification performance was significantly different (p < 0.05in a 2-tailed permutation test) from the performance of a model trained on randomly shuffled labels. B. Model accuracy for non-unique hues. The accuracy of the model for non-unique hues is shown in a manner analogous to A, with the orange and blue colours representing the orange and turquoise stimuli respectively. C. Misclassification probabilities. Given the EEG response (at a given time-point) to one of the four hues, the model can either make an accurate prediction of the label (panels A and B), or misclassify the input. Each of the four subpanels here shows the prediction probabilities for one particular input label (shown on the top-left, above each subpanel), thus corresponding to one row of the confusion matrix. For instance, the first subpanel shows the probabilities (at each time-point) that the model classifies EEG responses to unique red stimuli as being elicited by unique red (accuracy), unique green, orange or turquoise stimuli. The colour coding for the four stimulus hues in each subpanel is the same as panels A and B. Also see Supplementary Video V1, which shows how the confusion matrix changes as a function of time elapsed from stimulus onset.

deflection at about 120-220 ms after stimulus onset (Berninger et al., 1989; Murray et al., 1987; Tobimatsu et al., 1996), but this response is significantly altered by the addition of luminance contrast (Rabin et al., 1994). Furthermore, while observer isoluminance drives ERPs in a manner closely resembling nominal isoluminance, any substantial changes in luminance contrast have been found to result in highly dissimilar waveforms (Rabin et al., 1994). Xing et al. (2015) demonstrate that this is due to non-linear interactions between colour and luminance signals, which is likely to result from the involvement of colour-luminance neurons. To assess the impact of potential colour-luminance interactions on classifier performance, we constructed a model that evaluated how decoding performance was affected when the model was trained on inputs which differ not only in hue but also luminance-contrast. We trained tECOC classifiers for each observer using 12 labels, corresponding to three different luminance-contrast levels for each of the four hues. In Fig. 4, we present the performance of our model in a manner similar to Fig. 2C. Each panel is one row of the confusion matrix, i.e., given the EEG signals for an input stimulus, it shows the prediction probabilities for all 12 labels. The hue of the input is denoted by the row (labelled in the right margin) and its luminance-contrast by the column (labelled on top). The same colours as Fig. 2C are used to denote the four hues. In addition, for each hue, we also use two additional brightness levels to represent the two luminance contrast ratios (thus, for a given hue, isoluminant stimulus is the least bright, 45% luminance contrast is brighter, and 90% luminance contrast is the brightest). We find that while isoluminant signals can indeed be classified 100-300 ms after stimulus onset (left column), addition of luminance information disrupts the model performance for all hues (middle and right columns). Furthermore, we find that the classifier does not confuse isoluminant and non-isoluminant stimuli. This suggests that in contrast to a change in stimulus-shape where the temporal structure of hue-related information was preserved (Fig. 3), addition of luminance-contrast to the stimulus disrupts the temporal patterns which encode hue-information. The above observations also hold when separate models are trained for each luminance contrast condition (Supplementary Figure S5).

To characterise the effect of luminance, we trained a model using only the luminance labels of EEG signals (i.e., we used three labels corresponding to the three contrast levels). We found that all luminance



Fig. 3. Decoding performance for active and passive tasks is very similar, despite large differences in stimulus-evoked activity. A. Global field power. The left side of the panel shows the topographies and the Global Field Power (GFP) for stimulus onset. The hues are colour-coded (unique green is shown in green, red in red, etc.), and each panel shows the GFP for one luminance-contrast condition. The stimulus onset is marked by a dashed line at 0 abscissa. The right side of the panel shows the same for the shape-change event. B. Robustness to task. Separate models were trained using passive viewing and shape-change segments. Each subpanel shows the accuracy of the two models for on a particular input hue (e.g., the leftmost panel shows the model accuracy when EEG responses to red stimuli were used as inputs). The performance of the passive-segment model is shown using the same colours and symbols as Fig. 2A, while the shape-change model is shown using a dashed line for observer mean and darker shading for the bootstrapped 95% confidence intervals. Horizontal lines underneath show time-points where a randomised model trained on shuffled labels was significantly different from the model trained on correct labels.

conditions (Fig. 5A) can be decoded to above-chance levels, with the isoluminant and 90% contrast conditions being the most decodable. This is likely to reflect the fact that while both isoluminant and high-contrast stimuli are relatively easy to discriminate, the 45%-contrast stimuli are likely to contain characteristics resembling both these classes. An examination of the misclassification patterns of the model (Fig. 5B) further revealed that while isoluminant stimuli are robustly classified, the non-isoluminant conditions are more likely to be confounded with one another.

Stimuli with 45% luminance contrast have an above-chance probability of being misclassified as 90% luminance contrast. However, this effect is not symmetric, with 90% luminance contrast being easier to detect compared to the 45% contrast. Thus, under non-isoluminant conditions, not only are the hue-driven patterns difficult to detect, but they seem to be progressively overridden by luminance-contrast-driven patterns. To ensure that this effect was driven by luminance, and not by the chromatic content of the stimuli, we set up separate models for each hue, and were able to confirm that the effect was indeed independent of the stimulus hue. For each hue, the isoluminant stimuli were robustly classified (Supplementary **Figure S**6, leftmost column), while the non-isoluminant conditions produced similar but asymmetric prediction scores (Supplementary **Figure S**6, middle and right columns).

3.2. Interim discussion

Our findings are in line with Sutterer et al. (2021) who recently reported that both colour and luminance content can be successfully decoded from EEG signals. Hermann et al. (2022) investigated decoding of hue or luminance polarity from MEG signals and found that generalising luminance polarity across hue works better than generalising hue across polarity. This is consistent with our own findings that decoding of hue

is strongly affected by the addition of luminance contrast. Unlike these studies, where only stimuli that combine colour and luminance contrast were used, we also included stimuli that were isoluminant with the background. We found that decoding of hue from such nominally isoluminant stimuli is much more efficient. We also find an asymmetry in decoding unique and intermediate hues, with superior performance for unique hues.

Hermann and colleagues (2022) speculate that alignment with the daylight locus might represent an important determinant of colour decoding. They find that low-level, cone-opponent chromatic content impacts hue decoding, with differential cone-opponent inputs along the L-M and S-(L + M) mechanisms providing separable input into EEG signals that are being decoded. This is not surprising, as L-M and S-(L + M) cone-opponent signals combine differently with luminance information (e.g., Martinovic and Andersen, 2018). In our data, red and turquoise are closest to a cone opponent axis (i.e., L-M) and were also more distant to their neighbouring colours. Yet we find superior decoding for red and green. Despite being able to discard fully reductive cone-opponent input or colour-distance accounts of hue-decoding asymmetries, our experiment remains inconclusive as to the potential source of the observed effects.

The stimulus set in Experiment 1 was designed to investigate whether unique hues have more robust EEG representations. To achieve this, we chose unique and non-unique hues that were maximally distant in a perceptual space – red and green, orange and turquoise (see details of the stimulus set in Methods). As already reported by Rosenthal et al. (2021) and Hermann et al. (2022), inter-hue differences in decoding efficiency manifest even between such evenly spaced colours. Indeed, our findings confirm that the neurometric colour space is non-uniform, even when stimuli are made isoluminant; but they also suggest that unique hues may have a more distinct neural representa-



Fig. 4. Luminance information disrupts hue decoding. EEG responses to unique (green and red) and non-unique (orange and turquoise) hues at three luminancecontrasts (isoluminant, 45% and 90%) were used to train a tECOC model. Each of the 12 subpanels in this figure represents one row of the confusion matrix (similar to Fig. 2C). This corresponds to presenting the trained model with EEG responses to a given stimulus class, and observing the classification probabilities for all classes, including the input class. The hue and luminance contrast of the input labels are denoted by the row and column respectively. For each predicted label, the hue is represented by the corresponding colour (green, red, orange and turquoise), and the luminance-contrast by the brightness (isoluminant: lowest brightness, 45% contrast: intermediate brightness, 90% contrast: highest brightness). Shaded region around the curves shows bootstrapped 95% confidence intervals. Horizontal lines underneath show time-points where the accuracy (the correct classification of the input, i.e., only diagonal elements of the confusion matrix time-series) of a randomised model trained on shuffled labels was significantly different from the model trained on correct labels.

tion, indexed through superior information decoding when compared to intermediate hues. In our next experiment we aimed to further investigate the origin of the more robust decoding for unique as opposed to intermediate hues by introducing proximal neighbours, clockwise and counter clockwise to each hue. Decoding colours in such small neighbourhoods allows us to understand how perceptual notions of huedifference map to the EEG-derived neurometric space. If the decoding manifold contains three highly-proximal stimuli, which are only about 2 perceptual JNDs apart, this may result in a failure to decode if their EEG signatures are too similar. If, however, unique hues have a more robust EEG signature, they may be more decodable from their neighbours.

We also changed the four hues, replacing red with yellow. To disambiguate if unique or intermediate hue status drives a more robust neural signal irrespective of daylight locus alignment, which has been suggested as the source of asymmetries in the neurometric space by Hermann et al. (2022), it would be necessary to use a unique hue that is also more aligned with the daylight locus, such as yellow or blue.

3.3. Experiment 2: decoding over small and large perceptual hue differences

In Experiment 1 we showed superior decoding performance for unique hues compared to intermediate hues, suggesting a robust neural representation for the former. In Experiment 2, this hypothesis was further critically tested by using both small and large hue differences. Our aim was to re-examine decoding of nominally isoluminant unique and intermediate hues with a slightly modified hue set (see Interim Discussion above) and to extend it by decoding local clusters of stimuli around each of these hues. First, we measured individual settings for unique (yellow and green) and non-unique (orange and turquoise) hues for each observer. Next, we made EEG measurements in a task analogous to Experiment 1. For each observer, we used a stimulus set consisting of their individual settings for the four hues (denoted as the = configuration), and two sets of stimuli generated by rotating the individual settings by $\pm 10^{\circ}$ in CIELAB colour space (denoted as the + and – configurations respectively) – leading to a total of 12 stimuli (4 hue-clusters



Fig. 5. Luminance decoding from EEG signals. A. Mean classification accuracy. This panel shows the performance of the model in correctly identifying the luminance contrast of the stimuli (model accuracy). Each line shows the accuracy for one condition, with dark grey coding for the isoluminant condition, medium grey coding for 45% luminance contrast, and light grey coding for 90% luminance contrast (coding of luminance contrast using lightness is used throughout the article). B. Misclassification probabilities. Each subpanel shows one row of the confusion matrix analogous to Fig. 2C. The left panel shows classification probabilities for the three luminance conditions when isoluminant stimulus is presented to the classifier. Similarly, the middle and right panels show prediction probabilities when 45% and 90% luminance contrast inputs are presented to the classifier. In all panels, the shaded area around the lines shows bootstrapped 95% confidence intervals. Chance performance is shown by the dashed line. Horizontal lines underneath show time-points where a randomised model trained on shuffled labels was significantly different from the model trained on correct labels.

and 3 rotational-configurations, see Fig. 1C). The individual hue settings were distributed as follows (means and SEs): yellow $101^{\circ} \pm 2^{\circ}$, orange $61^{\circ} \pm 3^{\circ}$, green $153^{\circ} \pm 3^{\circ}$ and turquoise $198^{\circ} \pm 3^{\circ}$

In the shape discrimination task, grand mean accuracy was 96% \pm 1% SE (see Supplementary **Figure S7A**) and reaction times were 706 \pm 61 ms (See Supplementary **Figure S7B**). Response-time data were analysed with a 4 × 3 repeated measures ANOVA (4 hues vs. 3 rotational configurations, i.e., -, + and = sets), which yielded a significant main effect of hue (F (1.77, 26.5) = 5.25, *p*=.01, η p2 = 0.26) and an interaction with the rotational configuration (F (2.16, 32.49) = 5.08, *p* = 0.01, η p2 = 0.25) while the effect of the rotation itself was not significant (F (1.79, 26.99) = 0.72, *p* = .48, η p2 = 0.05). The interaction was deconstructed by separate repeated measures ANOVAs at each hue: for yellow,

there was a significant effect of rotation (F(1.36,20.48) = 6.23, p = 0.01, $\eta p = 0.29$) driven by slower RTs for the individual hue setting vs. 10° clockwise setting (p = .006). For green, there was also a significant effect (F(1.58,23.74) = 6.76, p = 0.007, $\eta p = 0.31$) driven by faster RTs for the individual hue setting vs. 10° clockwise setting (p = .04) as well as vs. 10° counterclockwise setting (p = .005); no differences were found for orange (p=.22) and for turquoise (p=.11). Taken together, we can see that only for unique hues (yellow and green) the responses to individual hue settings (= configuration) seem to be different from responses to $\pm 10^{\circ}$ rotated hues (i.e., – and + configurations). However, the direction of the effect was opposite for the two hues – while participant responded slower to their individual yellow setting, they responded faster to their individual green setting.



Fig. 6. Representational similarity between perceived hue and decoder performance. A. Dissimilarity in classifier outputs. tECOC models were trained to classify twelve colours from their EEG responses. The colours sampled four clusters along the hue circle corresponding to orange (O), yellow (Y), green (G), and turquoise (T), with each cluster consisting of settings made by the observer in a psychophysical experiment (=), and colours sampled 10° clockwise (-) and anti-clockwise (+) with respect to each setting. Each panel shows a dissimilarity matrix derived from classifier output. The panels show the dissimilarity 50, 150, 250, and 450 ms after stimulus onset. B. Mean classifier dissimilarity. The average dissimilarity over the period where the correlation is statistically significant. C. Dissimilarity in perceptual space. Hue angles of the 12 stimuli (same as panel A) were used to estimate dissimilarity in the perceptual CIELAB space. D. Representational similarity. Rank-correlation between perceptual and classification dissimilarities using Kendall's tau statistic. The solid curve shows the mean statistic, while the shaded envelope shows bootstrapped 95% confidence intervals over the observers. The horizontal line underneath shows time-points where the correlation was statistically significant (p < 0.05).

For the categorical rating task, the average ratings and their SEs were as follows: individual yellow 5.62 ± 0.6 ; -10° yellow 5.75 ± 0.57 ; $+10^{\circ}$ yellow 2.56 ± 0.35 ; individual green 6.93 ± 0.26 ; -10° green 7.93 ± 0.17 ; $+10^{\circ}$ green 3.87 ± 0.35 ; individual orange 6.37 ± 0.36 ; -10° orange 3.93 ± 0.48 ; $+10^{\circ}$ orange 7.25 ± 0.48 ; individual turquoise 5.68 ± 0.53 ; -10° turquoise 7.43 ± 0.53 ; $+10^{\circ}$ turquoise 3.18 ± 0.5 . We used these categorical ratings as a control measure, evaluating if the decoding performance can be reduced to proximity to focal colours. Again, no such relations were found (See Supplementary **Figure S7C**).

3.3.1. Decoding over large hue differences is predicted by hue angles

For each observer, we trained tECOC models over all stimuli: the four hue settings (= group), and the eight stimuli generated by $\pm 10^{\circ}$ rotations of each of these settings (+ and - groups respectively). Using the classification results, we generated a time-series of dissimilarity matrices (see Methods for details) and found that the stimulus representations were dissimilar in a 100–350 ms window after stimulus onset (Fig. 6A), with a stable mean dissimilarity (Fig. 6B). Similarly, we also calculated a perceptual dissimilarity measure by using differences in hue angles of the stimuli in CIELAB space. As expected, perceptual dissimilarity increases as one moves away from a given reference stimulus (Fig. 6C). Using rank-correlation analysis, we found the Kendall's tau statistic to be significant (p < 0.05) in a 100–350 ms range post-stimulus (Fig. 6D), suggesting that perceptual distances are correlated with decoding output.

3.3.2. Local distortions in hue decoding

Next, we posed the question: is the perceptual robustness of unique hues reflected in the structure of the decoding space around their respective representations? To answer this question, we trained 4 tECOC models – one on each of the four hue-clusters (orange, yellow, green, and turquoise). A hue cluster consisted of the observer's individual settings (=), and stimuli 10° clockwise (+) and counter-clockwise (-) from the individual settings. Each model was trained to classify EEG signals (responses to stimuli drawn from the respective cluster) into one of the

three labels =, +, or –. In Fig. 7A we show the results for the four models, one model per row. Each subpanel is a row in the corresponding confusion matrix, with the test stimulus indicated on top. For instance, the first row corresponds to the model trained on the yellow cluster. The first panel of this row shows the predictions of the model when stimuli 10° counter-clockwise from individual yellow settings were presented to it (i.e., the first row of the confusion matrix for the 'yellow' model).

We found that the three groups (individual settings, and the $\pm 10^{\circ}$ rotations) cannot be decoded in non-unique hues (Fig. 7A, first and fourth rows). However, for unique hues (Fig. 7A, second and third row), the rotated groups (first and third columns) can be decoded, while the individual settings (second column) cannot. This could reflect relative differences in visually evoked potential (Global Field Power: Fig. 7B, Eventrelated Potentials: Fig. 7C), and suggests that the representational space around unique hues is anisotropic (Fig. 8). Note that in the perceptually uniform CIELAB space the three groups, by design, had equivalent relative distributions (- and + were simply mean-shifted copies of =).

4. Discussion

Our first finding is that - under isoluminant conditions - EEG responses to the three tested unique hues show more distinct patterns compared to non-unique hues, and these patterns are stable during both passive viewing (Fig. 2) and active task-engagement (Fig. 3). We can also reach certain conclusions about the underlying neural processes from the time-course of decoding performance. While additional analysis (see Supplementary **Figure S**8) shows that the classification performance is unlikely to be driven by stimulus cone-contrast, a 100–300 ms decoding window is consistent with the idea that the performance of the model could be driven by both perceptual and post-perceptual contributions (Forder et al., 2017b). This is supported by the fact that the decoding performance steadily rises before peaking between 150 and 200 ms after stimulus onset, a time-window where EEG signals begin reflecting postvisual evaluative processing (VanRullen and Thorpe, 2001), including colour categorisation (Fonteneau and Davidoff, 2007). The chromatic



Fig. 7. Local distortions in representational space and Visually Evoked Potentials. A. Decoding within colour clusters. tECOC models were trained on clusters around the individual settings for unique (yellow: Y, green: G) and non-unique (Orange: O, Turquoise: T) hues. Each cluster consisted of three groups: individual observer settings (=), and two groups derived from 10° clockwise (+) and counter-clockwise (-) rotations of the individual settings in CIELAB space. Each row shows a model trained on a different hue (top row: orange, second row: yellow, etc.), with subpanels showing rows of the corresponding confusion matrices. E.g., the first row shows the classifier trained on the yellow cluster, and the first panel of this row shows the row of the confusion matrix that corresponds to the Y- input. The input stimulus for each row of the confusion matrix is labelled on top. B. Global Field Power. Each panel shows the mean Global Field Power (GFP) within a colour cluster (the name of the cluster is above each panel). The GFP for the central colour (=) in each cluster is shown in medium grey, while a darker and lighter grey are used to denote GFP for colours -10° (-) and $+10^{\circ}$ (+) to the central colour respectively. C. Event Related Potential. Mean Event Related Potentials calculated across observers. Only the occipital electrodes used for the decoding analysis are considered. The same colour scheme as B is used.



Fig. 8. Effect of representational anisotropies on decoding. The left panel shows a configuration of the distributions for the three groups (-, = and +) which can lead to poor decoding scores such as those observed within clusters of non-unique hues. The distributions overlap, and the distances between the distribution means $(\mathbf{r}^-, \mathbf{r}^=, \mathbf{r}^+)$ are too narrow to allow for proper discrimination using a linear boundary. The two other panels show possible representational anisotropies around the central hue which can lead to better discrimination in the neighbourhood. A. Scenario 1: Reduced variability. This panel shows how a relative decrease in the variability of unique hue representations (=) can lead to better decoding for the neighbouring hues (- and +). B. Scenario 2: Expansive anisotropy. This panel shows how a dilation of representational distances in the neighbourhood of unique hues (=) can also lead to increased decoding scores for the surrounding hues.

visual evoked potential (cVEP), which reflects the activation of colour sensitive neurons in visual cortices, also remains maximal in the same time window (Nunez et al., 2018). Due to the manner in which decoding from EEG signals is commonly implemented, the decoding performance is also reflective of the visual evoked potential patterns (Fig. 3A, and Fig. 7B and C). Since single trials are binned and averaged prior to decoding to increase signal-to-noise ratio (e.g., see Al-shargie et al., 2018; Bae and Luck, 2018 for similar ECOC-based approaches), the decoding necessarily reflects differences between dominant VEP components during the window of above-chance classification. In this study, these largely correspond to the P1/N1 window. The P1 component (peaking ~80-120 ms) is driven by luminance contrast and saturates once this contrast exceeds ~16% Weber contrast (Ellemberg et al., 2001; see also, Zemon and Gordon, 2006). Meanwhile, the N1 component (peaking ~150-200 ms) reflects both luminance and chromatic contrast and does not saturate (for normative data, see Porciatti and Sartucci, 1999). When luminance contrast is absent or low (i.e., at isoluminance as well as near it; see Rabin et al., 1994) it is the chromatic signals that drive the VEPs, influencing both their peak latency and amplitude in ways that would create more distinctive response representations across hues. Pitzalis et al. (2018) performed fMRI-guided source analysis of the cVEP and found that between 100 and 300 ms it was mainly driven by V1 and V8/VO activity, including feedforward and recurrent connections between them and other colour-sensitive areas. However, a high-level interpretation of the decoding on the basis of the categorical status of the stimulus colours is unlikely. Categorical representativeness ratings do not follow the pattern observed in the classifier performance (see Supplementary Figure S3C). The most parsimonious explanation for the pivots in colour space that drive asymmetries in decoding around unique hue locations would be that they correspond to hue locations that are associated with a more robust neural representation, thus making it more easily decodable from less robustly represented hues.

Secondly, classification performance for the decoding of hues diminished when luminance contrast was added (Fig. 4 and Supplementary Figure S5). This was not entirely unexpected since luminance contrast is known to have a strong effect on EEG responses, once luminance contrast is sufficiently strong (Rabin et al., 1994). At the same time, we found that hues could be decoded at all luminance levels at abovechance levels within the same 100–300 ms window (Fig. 5 and Supplementary Figure S2). Thus, under non-isoluminant conditions, not only are the hue-driven patterns more difficult to detect, but they may also be at least partly overridden or replaced by luminance-contrast or joint-colour-and-luminance-contrast-driven activity. Our findings are consistent with the idea that in the visual cortex, hue is most likely to be encoded by neural populations which also encode luminance (see also Conway et al., 2007 for work in the extrastriate cortex). The fact that purely chromatic-tuned cells in the visual cortex are known to be in a minority compared to luminance-tuned or luminance-chromaticity tuned cells (Johnson et al., 2001; Lennie et al., 1990) may partly explain why luminance signals tend to override chromatic information in EEG recordings from the occipital cluster. In V1-V3, the neurons are tuned to many intermediate directions, both in terms of hue and luminance contrast (for a review, see Gegenfurtner and Kiper, 2003). In higher-level areas of the extrastriate cortex, colour representations become organised in ways that resemble perceptual colour spaces (for macaque neurophysiology, see Bohon et al., 2016; Conway et al., 2007; and for fMRI in humans, see Brouwer and Heeger, 2013, 2009). Thus, the decoding in our study is likely to reflect cumulative effects that build up across these areas. Even though we find more robust responses for the two unique hues (red and green) compared to the two non-unique hues (orange and turquoise), decoding is still possible for non-unique hues, implying that there are indeed multiple hue representations that are being encoded by the brain (see, e.g., Brouwer and Heeger, 2009; Parkes et al., 2009; Zaidi and Conway, 2019).

Thirdly, we show in Experiment 2, that the geometric structure of this representational space can be explored by carefully designed experiments. Our results demonstrate that while large distances in the neural representational space are indeed correlated with perceptual hue differences (Fig. 6), there are local anisotropies associated with unique hues (Fig. 7) which are likely to represent local changes in representational geometry and variability. Fig. 8 illustrates two such scenarios which may be particularly relevant to unique hues. One possibility is the narrowing of representational probabilities around unique hues (middle panel, Fig. 8). This would correlate with the reduced variabilities reported in large datasets of unique hue measurements (Xiao et al., 2013), and suggest that the reduced perceptual variability is also reflected in reduced variability in cortical representations. A second possibility is the dilation of the representational space in the neighbourhood of unique hues (right pane, Fig. 8). This scenario would imply that there is an increase in the number of possible hue representations in the neighbourhood of unique hues compared to non-unique hues - reflecting an increase in cortical resources used for encoding. Note that these two scenarios are not exclusive, and by no means unique. Such tunings could reflect properties of our environment such as the statistical regularities in the reflectance spectra of naturally occurring surfaces (Philipona and O'Regan, 2006) or the degree of alignment with the daylight locus (Hermann et al., 2022). Alternatively, they may also be, at least in part, driven by a broader categorical distinction between warm and cool colours (Rosenthal et al., 2021). Perhaps this is the reason why the neural reality of perceptual red-green and blue-yellow hue-opponent mechanisms has proven to be so elusive - it is not a fundamental mechanism hard-wired into the neural circuitry, but a statistical peak in the tuning of neural populations, many of which multiplex both colour and luminance information. Such complex influences on the neural encoding would make hue-specific signals much harder to detect.

A growing number of studies investigating population activity analyse EEG and MEG topographical data by interrogating timepoint-bytimepoint trajectories in activation manifolds. Our results suggest that the structure of such manifolds can be highly anisotropic, and that these anisotropies can reflect perceptual measurables. In the case of hue perception, it is possible that the local structure of this space is reflected in quasi-invariants such as the so-called unique hue percepts. Several reports have already established that a neurometric mapping of hue spaces using EEG information decoding is viable (Hajonides et al., 2021; Hermann et al., 2022; Rosenthal et al., 2021). This study marks a first hypothesis-based exploration of these maps and shows that unique hues may represent local anisotropies in cortical hue-representations.

Competing Interests

The authors declare no competing interests.

Credit authorship contribution statement

Tushar Chauhan: Conceptualization, Methodology, Software, Validation, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **Ivana Jakovljev:** Investigation, Data curation, Validation. **Lindsay N. Thompson:** Investigation, Data curation, Validation. **Sophie M. Wuerger:** Conceptualization, Writing – original draft, Writing – review & editing. **Jasna Martinovic:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing. Jasna Martinovic: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Data availability

tECOC has been released as open source software. Link to the data has also been provided.

Acknowledgments

TC was partially supported by the grant DE-SC0022997 (US Department of Energy), and a Picower Fellowship from The JPB Foundation. JM was partly supported by grant BB/R009287/1 from the Biotechnology and Biological Sciences Research Council (BBSRC). IJ was supported by a scholarship for postdoctoral training of researchers awarded by the Ministry of Education, Science and Technological Development, Republic of Serbia.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.119884.

References

- Al-shargie, F., Tang, T.B., Badruddin, N., Kiguchi, M., 2018. Towards multilevel mental stress assessment using SVM with ECOC: an EEG approach. Med. Biol. Eng. Comput. 56, 125–136. doi:10.1007/s11517-017-1733-8.
- Bae, G.-Y., Luck, S.J., 2018. Dissociable Decoding of Spatial Attention and Working Memory from EEG Oscillations and Sustained Potentials. J. Neurosci. 38, 409–422. doi:10.1523/JNEUROSCI.2860-17.2017.
- Berninger, T.A., Arden, G.B., Hogg, C.R., Frumkes, T., 1989. Separable evoked retinal and cortical potentials from each major visual pathway: preliminary results. British J. Ophthalmol. 73, 502–511. doi:10.1136/bjo.73.7.502.
- Bohon, K.S., Hermann, K.L., Hansen, T., Conway, B.R., 2016. Representation of perceptual color space in macaque posterior inferior temporal cortex (the V4 Complex). eNeuro 3. doi:10.1523/ENEURO.0039-16.2016.

Brouwer, G.J., Heeger, D.J., 2013. Categorical clustering of the neural representation of color. J. Neurosci. 33, 15454–15465. doi:10.1523/JNEUROSCI.2472-13.2013.

Brouwer, G.J., Heeger, D.J., 2009. Decoding and reconstructing color from responses in human visual cortex. J. Neurosci. 29, 13992 LP –14003. CIE, 2004. 15.3: 2004 Colorimetry, Vienna, Austria: CIE Central Bureau. CIE.

- Clifford, A., Holmes, A., Davies, I., Franklin, A., 2010. Color categories affect pre-attentive color perception. Biol. Psychol. 85, 275–282. doi:10.1016/j.biopsycho.2010.07.014. Conway, B., Stoughton, C., 2009. Response: towards a neural representation for unique
- hues. Current Biol. 19, R442–R443. doi:10.1016/j.cub.2009.04.056. Conway, B.R., Moeller, S., Tsao, D.Y., 2007. Specialized color modules in Macaque Extras-
- triate Cortex. Neuron 56, 560–573. doi:10.1016/j.neuron.2007.10.008. De Valois, R., De Valois, K., 1993. A multi-stage color model. Vision Res. 33, 1053–1065.
- doi:10.1016/042-6989(93)9024 P.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009.
- Ellemberg, D., Hammarrenger, B., Lepore, F., Roy, M.-.S., Guillemot, J.-.P., 2001. Contrast dependency of VEPs as a function of spatial frequency: the parvocellular and magnocellular contributions to human VEPs. Spat. Vis. 15, 99–111. doi:10.1163/15685680152692042.
- Fairchild, M., 2013. Color Appearance Models. John Wiley & Sons, Ltd, Chichester, UK doi:10.1002/9781118653128.
- Fonteneau, E., Davidoff, J., 2007. Neural correlates of colour categories. Neuroreport 18, 1323. doi:10.1097/WNR.0b013e3282c48c33.
- Forder, L., Bosten, J., He, X., Franklin, A., 2017a. A neural signature of the unique hues. Sci. Rep. 7, 42364. doi:10.1038/srep42364.
- Forder, L., He, X., Franklin, A., 2017b. Colour categories are reflected in sensory stages of colour perception when stimulus issues are resolved. PLoS ONE 12, e0178097. doi:10.1371/journal.pone.0178097.
- Gegenfurtner, K.R., Kiper, D.C., 2003. Color Vision. Annu. Rev. Neurosci. 26, 181–206. doi:10.1146/annurev.neuro.26.041002.131116.
- Goddard, E., Mullen, K., 2020. fMRI representational similarity analysis reveals graded preferences for chromatic and achromatic stimulus contrast across human visual cortex. Neuroimage 215, 116780. doi:10.1016/j.neuroimage.2020.116780.
- Hajonides, J., Nobre, A., van Ede, F., Stokes, M., 2021. Decoding visual colour from scalp electroencephalography measurements. Neuroimage 118030. doi:10.1016/j.neuroimage.2021.118030.
- Hering, E., 1920. Outlines of a Theory of the Light Sense. Translated from German by L. Hurvich & D. Jameson 1964. Harvard Univ. Press, Cambridge.
- Hermann, K., Singh, S., Rosenthal, I., Pantazis, D., Conway, B., 2022. Temporal dynamics of the neural representation of hue and luminance polarity. Nat Commun 13, 661. doi:10.1038/s41467-022-28249-0.
- Hinks, D., Cárdenas, L., Kuehni, R., Shamey, R., 2007. Unique-hue stimulus selection using Munsell color chips. J. Optical Soc. Am. A 24, 3371. doi:10.1364/JOSAA.24.003371.
- Holmes, A., Franklin, A., Clifford, A., Davies, I., 2009. Neurophysiological evidence for categorical perception of color. Brain Cogn 69, 426–434. doi:10.1016/j.bandc.2008.09.003.
- Jameson, D., Hurvich, L., 1964. Theory of brightness and color contrast in human vision. Vision Res. 4, 135–154. doi:10.1016/0042-6989(64)90037-9.
- Jazayeri, M., Afraz, A., 2017. Navigating the Neural Space in Search of the Neural Code. Neuron 93, 1003–1014. doi:10.1016/j.neuron.2017.02.019.
- Johnson, E., Hawken, M., Shapley, R., 2001. The spatial transformation of color in the primary visual cortex of the macaque monkey. Nat. Neurosci. 4, 409–416. doi:10.1038/86061.
- Johnson, K., 2000. Neural Coding. Neuron 26, 563–566. doi:10.1016/S0896-6273(00)81193-9.
- Junghöfer, M., Elbert, T., Tucker, D.M., Rockstroh, B., 2000. Statistical control of artifacts in dense array EEG/MEG studies. Psychophysiology 37, 523–532. doi:10.1111/1469-89866.3740523.
- Kaneshiro, B., Guimaraes, M.P., Kim, H.-.S., Norcia, A.M., Suppes, P., 2015. A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification. PLoS ONE 10, e0135697. doi:10.1371/journal.pone.0135697.
- Knoblauch, K., Bieber, M.L., Werner, J.S., 1998. M- and L-cones in early infancy: I. VEP responses to receptor-isolating stimuli at 4- and 8-weeks of age. Vision Res. 38, 1753– 1764. doi:10.1016/S0042-6989(97)00383-0.
- Lee, B.B., Martin, P.R., Grünert, U., 2010. Retinal connectivity and primate vision. Prog. Retin. Eye Res. 29, 622–639. doi:10.1016/j.preteyeres.2010.08.004.
- Lennie, P., Krauskopf, J., Sclar, G., 1990. Chromatic mechanisms in striate cortex of macaque. J. Neurosci. 10, 649–669. doi:10.1523/JNEUROSCI.10-02-00649.1990.
- Maris, E., 2012. Statistical testing in electrophysiological studies. Psychophysiology 49, 549–565. doi:10.1111/j.1469-8986.2011.01320.x.
- Martinovic, J., Andersen, S., 2018. Cortical summation and attentional modulation of combined chromatic and luminance signals. Neuroimage 176, 390–403. doi:10.1016/j.neuroimage.2018.04.066.
- Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48, 229–240. doi:10.1111/j.1469-8986.2010.01061.x.
- Mollon, J., 2009. A neural basis for unique hues? Curr. Biol. 19, R441–R442. doi:10.1016/j.cub.2009.05.008, author reply R442-3.
- Murray, I., Parry, N., Carden, D., Kulikowski, J., 1987. Human visual evoked potentials to chromatic and achromatic gratings. Clin. Vis. Sci. 1, 231–244.
- Namima, T., Yasuda, M., Banno, T., Okazawa, G., Komatsu, H., 2014. Effects of luminance contrast on the color selectivity of neurons in the macaque area V4 and Inferior Temporal Cortex. J. Neurosci. 34, 14934–14947. doi:10.1523/JNEUROSCI.2289-14.2014.
- Nolan, H., Whelan, R., Reilly, R.B., 2010. FASTER: fully automated statistical thresholding for EEG artifact rejection. J. Neurosci. Methods 192, 152–162. doi:10.1016/j.jneumeth.2010.07.015.
- Nunez, V., Shapley, R., Gordon, J., 2018. Cortical double-opponent cells in color perception: perceptual Scaling and chromatic visual evoked potentials. Iperception 9. doi:10.1177/2041669517752715.

- Nunez, V., Shapley, R., Gordon, J., 2017. Nonlinear dynamics of cortical responses to color in the human cVEP. J. Vis. 17, 9. doi:10.1167/17.11.9.
- Parkes, L.M., Marsman, J.-B.C., Oxley, D.C., Goulermas, J.Y., Wuerger, S.M., 2009. Multivoxel fMRI analysis of color tuning in human primary visual cortex. J Vis 9. doi:10.1167/9.1.1, 1–1.
- Philipona, D., O'Regan, K., 2006. Color naming, unique hues, and hue cancellation predicted from singularities in reflection properties. Vis. Neurosci. 23, 331–339. doi:10.1017/S0952523806233182.
- Pitzalis, S., Strappini, F., Bultrini, A., Di Russo, F., 2018. Detailed spatiotemporal brain mapping of chromatic vision combining high-resolution VEP with fMRI and retinotopy. Hum. Brain Mapp. 39, 2868–2886. doi:10.1002/hbm.24046.
- Porciatti, V., Sartucci, F., 1999. Normative data for onset VEPs to red-green and blue-yellow chromatic contrast. Clin. Neurophysiol. 110, 772–781. doi:10.1016/S1388-2457(99)00007-3.
- Rabin, J., Switkes, E., Crognale, M., Schneck, M., Adams, A., 1994. Visual evoked potentials in three-dimensional color space: correlates of spatio-chromatic processing. Vision Res. 34, 2657–2671. doi:10.1016/0042-6989(94)90222-4.
- Regan, B., Reffin, J., Mollon, J., 1994. Luminance noise and the rapid-determination of discrimination ellipses in color deficiency. Vision Res. 34, 1279–1299. doi:10.1016/0042-6989(94)90203-8.
- Ritchie, J.B., Kaplan, D.M., Klein, C., 2019. Decoding the brain: neural representation and the limits of multivariate pattern analysis in cognitive neuroscience. Br J Philos Sci 70, 581–607. doi:10.1093/bjps/axx023.
- Rosenthal, I., Singh, S., Hermann, K., Pantazis, D., Conway, B., 2021. Color Space Geometry uncovered with magnetoencephalography. Curr. Biol. 31, 515–526. doi:10.1016/j.cub.2020.10.062.
- Schanda, J., 2016. CIE u', v' uniform chromaticity scale diagram and CIELUV color space. Encyclopedia of Color Science and Technology. Springer-Verlag, New York, New York, USA.
- Shepard, R.N., 1958. Stimulus and response generalization: deduction of the generalization gradient from a trace model. Psychol. Rev. 65, 242–256. doi:10.1037/h0043083.
 Smith, T., Guild, J., 1931. The C.I.E. colorimetric standards and their use. Transactions of
- the Optical Society 33 (3), 72. doi:10.1088/1475-4878/33/3/01. Stoughton, C., Conway, B., 2008. Neural basis for unique hues. Curr. Biol. 18, R698–R699.
- doi:10.1016/j.cub.2008.06.018.

- Sutterer, D., Coia, A., Sun, V., Shevell, S., Awh, E., 2021. Decoding chromaticity and luminance from patterns of EEG activity. Psychophysiology 58, e13779. doi:10.1111/psyp.13779.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J.-.R., 2009. Unconscious effects of language-specific terminology on preattentive color perception. PNAS 106, 4567–4570. doi:10.1073/pnas.0811155106.
- Tobimatsu, S., Tomoda, H., Kato, M., 1996. Human VEPs to isoluminant chromatic and achromatic sinusoidal gratings: separation of parvocellular components. Brain Topogr 8, 241–243. doi:10.1007/BF01184777.
- VanRullen, R., Thorpe, S., 2001. The Time course of visual processing: from early perception to decision-making. J. Cogn. Neurosci. 13, 454–461. doi:10.1162/08989290152001880.

Walsh, J., 1958. Photometry. Dover Publications.

- Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data – a practical approach. J. Neurosci. Methods, Cutting-edge EEG Methods 250, 34–46. doi:10.1016/j.jneumeth.2014.08.002.
- Wuerger, S., Atkinson, P., Cropper, S., 2005. The cone inputs to the unique-hue mechanisms. Vision Res. 45, 3210–3223. doi:10.1016/j.visres.2005.06.016.
- Wuerger, S., Xiao, K., 2015. Color vision, opponent theory. In: Encyclopedia of Color Science and Technology. Springer, pp. 1–6.
- Xiao, K., Fu, C., Mylonas, D., Karatzas, D., Wuerger, S., 2013. Unique hue data for colour appearance models. Part II: chromatic adaptation transform. Color Res. Appl. 38, 22– 29. doi:10.1002/col.20725.
- Xiao, K., Wuerger, S., Fu, C., Karatzas, D., 2011. Unique hue data for colour appearance models. Part I: loci of unique hues and hue uniformity. Color Res. Appl. 36, 316–323. doi:10.1002/col.20637.
- Xing, D., Ouni, A., Chen, S., Sahmoud, H., Gordon, J., Shapley, R., 2015. Brightnesscolor interactions in human early visual cortex. J. Neurosci. 35, 2226–2232. doi:10.1523/JNEUROSCI.3740-14.2015.
- Zaidi, Q., Conway, B., 2019. Steps towards neural decoding of colors. Current opinion in behavioral sciences. Visual perception 30, 169–177. doi:10.1016/j.cobeha.2019.10.011.
- Zemon, V., Gordon, J., 2006. Luminance-contrast mechanisms in humans: visual evoked potentials and a nonlinear model. Vision Res. 46, 4163–4180. doi:10.1016/j.visres.2006.07.007.