Received: 12 May 2020 | Revised: 8 January 2021 | Accepted: 11 January 2021

DOI: 10.1111/psyp.13779

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Decoding chromaticity and luminance from patterns of EEG activity

David W. Sutterer[1,2,3] | **Andrew J. Coia**[1,2] | **Vincent Sun**[4] | **Steven K. Shevell**[1,2,5] | **Edward Awh**[1,2]

[1]Department of Psychology, University of Chicago, Chicago, IL, USA

[2]Institute for Mind and Biology, University of Chicago, Chicago, IL, USA

[3]Department of Psychology, Vanderbilt University, Nashville, TN, USA

[4]Center for Visual Communication and Color Research, Chinese Culture University, Taipei, Taiwan

[5]Department of Ophthalmology and Visual Science, University of Chicago, Chicago, IL, USA

Correspondence
David W. Sutterer, Department of Psychology, Vanderbilt University, 111 21st St, Nashville, TN 37212, USA.
Email: david.w.sutterer@vanderbilt.edu

Funding information
National Institute of Mental Health, Grant/Award Number: RO1 MH087214-08

Abstract

A long-standing question in the field of vision research is whether scalp-recorded EEG activity contains sufficient information to identify stimulus chromaticity. Recent multivariate work suggests that it is possible to decode which chromaticity an observer is viewing from the multielectrode pattern of EEG activity. There is debate, however, about whether the claimed effects of stimulus chromaticity on visual evoked potentials (VEPs) are instead caused by unequal stimulus luminances, which are *a*chromatic differences. Here, we tested whether stimulus chromaticity could be decoded when potential confounds with luminance were minimized by (1) equating chromatic stimuli in luminance using heterochromatic flicker photometry for each observer and (2) independently varying the chromaticity and luminance of target stimuli, enabling us to test whether the pattern for a given chromaticity generalized across wide variations in luminance. We also tested whether luminance variations can be decoded from the topography of voltage across the scalp. In Experiment 1, we presented two chromaticities (appearing red and green) at three luminance levels during separate trials. In Experiment 2, we presented four chromaticities (appearing red, orange, yellow, and green) at two luminance levels. Using a pattern classifier and the multielectrode pattern of EEG activity, we were able to accurately decode the chromaticity and luminance level of each stimulus. Furthermore, we were able to decode stimulus chromaticity when we trained the classifier on chromaticities presented at one luminance level and tested at a different luminance level. Thus, EEG topography contains robust information regarding stimulus chromaticity, despite large variations in stimulus luminance.

KEYWORDS

chromaticity, color vision, multivariate pattern analysis, visual evoked potential

1 | INTRODUCTION

Color vision has been studied for decades using electroencephalography (EEG). Discriminable visual evoked potentials (VEPs), typically recorded over the posterior occipital midline of the scalp (Murray et al., 1987; Paulus et al., 1984, 1986; Rabin et al., 1994; Skiba et al., 2014), are found when an observer views different chromaticities or luminances. VEP waveforms can be elicited by an equiluminant stimulus with chromatic variation, such as a stimulus that has sub-areas at the identical luminance but with different ratios of L-to-M cone activity (thus appearing, e.g., red and

green). These chromatic visual evoked potentials (cVEPs) are valuable tools used in both basic science (Nunez et al., 2018; Rabin et al., 1994) and clinical applications (Crognale, 2002; Crognale et al., 1993; Regan & Spekreuse, 1974). The chromatic response is thought to be separable from the waveform elicited by achromatic luminance, which sums L- and M-cone activity (Rabin et al., 1994). Also, work comparing waveforms at individual electrodes has shown that changes as a function of luminance contrast depend on the specific chromaticity of the stimulus (Klistorner et al., 1998), thus demonstrating that the waveform depends on chromaticity. Nevertheless, an open question is whether EEG activity contains sufficient information to discriminate specific chromaticity values (e.g., the chromaticity that appears "red" from the chromaticity that appears "green") in the absence of signal changes modulated by varying luminance. This question is addressed here.

Neuroimaging work has demonstrated that the specific stimulus chromaticity can be decoded from patterns of fMRI BOLD activity in visual cortex (Brouwer & Heeger, 2009; 2013), and there is broad interest in determining whether this information can be decoded from EEG activity. Recent work in the fields of working memory (Bocincova & Johnson, 2019) and brain-computer interface (BCI) development (Rasheed & Marini, 2015) aimed to decode the color of a stimulus from patterns of low frequency EEG activity on the scalp. VEPs, however, are well known to respond to differences in either stimulus chromaticity or luminance (Kulikowski et al., 1996; Skiba et al., 2014) and previous EEG and magnetoencephalography (MEG) work did not control for individual differences in luminance for the chromatic stimuli (Bocincova & Johnson, 2019; Rasheed & Marini, 2015; Sandhaeger et al., 2019). Furthermore, an important test for a true chromatic signature is whether classification of chromaticity is maintained despite changes in the luminance of the chromatic stimuli.

The work here seeks to test whether classification of chromaticity in human EEG is maintained despite changes in the luminance of chromatic stimuli, and also whether luminance can be decoded despite differences in chromaticity. Observers in two experiments monitored centrally presented chromatic disks while EEG was recorded. Two complementary techniques were used to rule out the possibility that luminance differences between stimuli could contaminate the chromatic signal. First, the luminance of chromatic stimuli at each luminance level was equated for each observer via heterochromatic flicker photometry ("HFP"; Lee et al., 1988). Second, both the luminance and chromaticity of a centrally presented disk were varied systematically, which allowed for training of a pattern classifier on the chromaticities presented at one luminance level and testing at the untrained luminance level. In the first experiment, one of two chromaticities (appearing red or green) at one of three luminance levels were presented on each trial. In the second experiment, four chromaticities (appearing red, orange, yellow, or green) were tested at two luminance levels.

To preview the results, the chromaticity of observed stimuli was successfully decoded using a pattern classifier trained on the topography of scalp EEG activity even after luminance differences were controlled using HFP. In addition, luminance levels of chromatic stimuli could be decoded as well. Furthermore, the chromaticity of the stimuli could be decoded by training the classifier on the chromaticities presented at one luminance level and then, testing at a different luminance level. Thus, multivariate analysis reveals a topographic signature of a given stimulus chromaticity that generalizes across wide variations in luminance.

2 | MATERIALS AND METHODS

2.1 | Observers

Sixteen volunteers (5 in Experiment 1, and 11 in Experiment 2) participated in the experiments for monetary compensation ($15/hr). Observers (7 female) were between 18 and 35 years old (M = 23.6, SD = 3.8), reported normal or corrected-to-normal visual acuity, and provided informed consent according to procedures approved by the University of Chicago IRB. Observers were screened for normal color vision using Ishihara color plates. Two observers' data were excluded from Experiment 2. One observer was excluded because of a technical problem during data collection and the other observer was excluded for failing to complete the behavioral task. Therefore, nine observers were included in the final sample for Experiment 2.

2.2 | Apparatus

Observers were tested in a dark, electrically shielded chamber. Stimuli were generated using MATLAB (Mathworks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and were presented on a 24" LCD monitor (BenQ XL2430T; 120 Hz refresh rate). Luminance levels of the background and stimuli were measured with a Photoresearch PR670 spectroradiometer.

2.3 | Heterochromatic flicker photometry in Experiment 1

A warm up routine was run for approximately 1 hour before the experiment began to stabilize the luminance output on the

monitor. RGB values corresponding to the desired luminance of the 5 cd/m^2 gray background (used in EEG portion) and "red" luminance levels were measured with a photometer before each experiment.

HFP stimuli consisted of a centrally fixated disk subtending 2.5° of visual angle. The chromaticity of the disk was exchanged in time between two chromaticities at a rate of 12.5 Hz (40 ms/chromaticity). Luminance was adjusted along only the R and G outputs of the display because the sum of the display's max R, G, and B luminances failed to add up to the max white luminance (a typical additivity failure of LCD displays). Adjusting the R and G outputs allowed for changing luminance while holding chromaticity constant. During HFP, disks were presented on a dark background (<0.1 cd/m^2). In separate trials the light that appeared red (hence called "red") was held constant at one of the three different luminance levels (6.5, 10.8, or 17.4 cd/m^2) while the light that appeared green ("green") was adjusted in luminance (100 steps from 0 to 50 cd/m^2). CIE (x, y) coordinates for the "red" and "green" stimuli were (0.61, 0.32) and (0.29, 0.55), respectively, (in CIELUV (u′, v′) coordinates, the "red" and "green" stimuli were (0.43,0.51) and (0.13,0.55), respectively).

Observers viewed the computer monitor at a distance of 74 cm while head position was stabilized with a chinrest. On each trial, observers viewed alternating "red" and "green" flickering disks and were instructed to adjust via keyboard button presses the level of the "green" (the left-arrow key raised the light level of the "green" disk and the right-arrow key lowered it) until the percept of flicker was minimized, at which point they would press a third button to make a match and end the trial.

The average of nine minimum flicker matches (three blocks each consisting of three trials) for each of the three "red" light levels used in the study was taken to be the equiluminant value for that observer for "green" stimuli used in the EEG portion of the experiment. The total time, including some practice trials to familiarize the observer with the task, was about 1 hour.

2.4 | Heterochromatic flicker photometry in Experiment 2

In Experiment 2, HFP was used to equate the luminance of chromaticities that appeared red, orange, yellow, or green, at two luminance levels (6.5 and 10.8 cd/m^2). First, the red/green match was measured as in Experiment 1, by holding the "red" disk at constant luminance and changing the "green" disk value. Next, in separate trials, "yellow" (CIE (x, y) = 0.39, 0.47; CIELUV (u′, v′) = 0.20,0.54), or "orange" (0.52,0.37; CIELUV (u′, v′) = 0.33,0.52) replaced "green," and then, "red" was adjusted by the observer for minimal-flicker. Based on this minimal-flicker "red" setting, the luminance of "yellow" or "orange" was then adjusted by the experimenter for the next block until the observer consistently selected a minimal-flicker "red" setting of 6.5 or 10.8 cd/m^2, as required for the experiment. "Yellow" luminance was adjusted by changing the red and green output values equally, while "orange" luminance was adjusted using a 2:1 red:green ratio. This allowed luminance changes of "yellow" or "orange" while chromaticity remained constant.

While the red/green matches at each luminance level were repeated for three blocks each, the red/orange and red/yellow conditions required extra blocks to reach the point at which an observer could consistently produce the specific "red" match value over three consecutive blocks. All values were set as the average of the final three blocks (nine trials) of minimum flicker match. The total time, including some practice trials to familiarize the observer with the task, was about 1 hr.

2.5 | EEG task procedure

In both experiments, observers performed a detection task in which they monitored the presentation of centrally presented chromatic disks on a gray background (CIE(x, y) = (0.31,0.32); 5 cd/m^2; CIELUV (u′, v′) = 0.20,0.46) for instances in which the disk was presented for longer than 100 ms (Figure 1). Observers initiated each trial by pressing the spacebar. Each trial began with a fixation point (0.2° in diameter) presented for a random duration between 500 and 800 ms. Next a colored disk (2.5° in diameter) was presented for either 100 ms (95% of trials) or for 1,500 ms (5% of trials). On short stimulus presentation (100 ms) trials, the colored disk was followed immediately by a 750 ms blank screen. On long stimulus presentation trials, observers were instructed to press the "?" key whenever they detected that the trial was longer than usual. On these long presentation trials, the stimulus was presented until the observer pressed the "?" key or 1,500 ms passed, whichever occurred first. If observers false alarmed and pressed the "?" key on short stimulus presentation trials, the trial was immediately aborted. The goal of this task was to ensure that observers paid attention to the stimuli. Thus, all long presentation trials and false alarm short presentation trials (maximum of seven trials for any participant) were discarded from EEG analysis. This task was used in both in Experiment 1 (Figure 1a) and Experiment 2 (Figure 1b) to ensure that observers were attending to the stimulus.

2.6 | Electrophysiology

In both Experiments, EEG was recorded from 30 active Ag/AgCl electrodes (Brain Products actiCHamp, Munich,

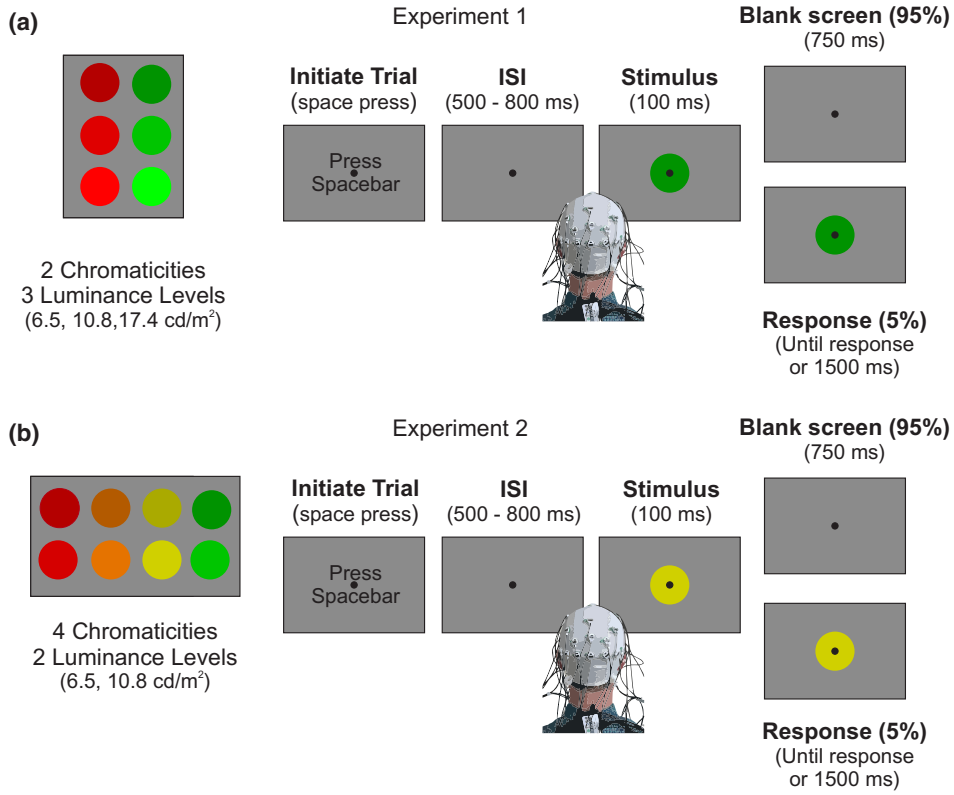

FIGURE 1 Task figure for Experiments 1 and 2. (a) Schematic of Experiment 1. (b) Schematic of Experiment 2

Germany) mounted in an elastic cap positioned according to the International 10–20 system Fp1, Fp2, F7, F3, F4, F8, Fz, FC5, FC6, FC1, FC2, C3, C4, Cz, CP5, CP6, CP1, CP2, P7, P8, P3, P4, Pz, PO7, PO8, PO3, PO4, O1, O2, and Oz, with a ground electrode at position FPz. Data were referenced online to the right mastoid and re-referenced offline to the algebraic average of the left and right mastoids. Incoming data were filtered (low cutoff = 0.01 Hz, high cutoff = 80 Hz, and slope from low to high cutoff = 12 dB/octave) and recorded with a 500 Hz sampling rate using Brain Vision Recorder (Brain Products, Munich, Germany) running on a PC. Selecting a high-pass filter cutoff that is too high (>0.5 Hz) can introduce substantial distortion in the data (Cohen, 2014; Luck, 2014) so we selected a low cutoff of 0.01 Hz for filtering the incoming data as it facilitates accurate artifact rejection by removing slow drifts while introducing minimal data distortion. Furthermore, this modest filtering during data acquisition had negligible effects on subsequent time–frequency decomposition (Cohen, 2014; Luck, 2014). EEG data were baselined over the 300 ms before disk onset. We recorded electrooculogram (EOG) with passive Ag/AgCl electrodes, which we used to monitor for eye movements and blinks. We recorded horizontal EOG from a bipolar pair of electrodes affixed ~1 cm from the external canthus of each eye, and recorded vertical EOG from a bipolar pair of electrodes affixed above and below the right eye.

2.7 | Eye tracking

Gaze position was monitored at a sampling rate of 500 Hz using a desk-mounted EyeLink 1000 Plus infrared eye-tracking camera (SR Research, Ontario, Canada) while head location was stabilized with a chin rest. Useable eye-tracking data were obtained for 4 of 5 participants in Experiment 1, and 10 of 11 participants in Experiment 2.

2.8 | Artifact rejection

Segmented EEG data were visually inspected for artifacts (amplifier saturation, excessive muscle noise, and skin potentials), and EOGs and gaze data for ocular artifacts (blinks and eye movements). Trials contaminated by artifacts were discarded. After artifact rejection, there were on average 898 trials per participant in Experiment 1 (*SD* = 172), and 1,410 trials per participant in Experiment 2 (*SD* = 145.8). In

Experiment 1, the average minimum number of trials in any color and luminance category (e.g., low luminance green) was 143 trials (*SD* = 29.3). In Experiment 2, the average minimum number of trials in any condition was 167 (*SD* = 19.6).

2.9 | Time–frequency analysis

To calculate frequency-specific activity at each electrode, we followed the same analytic procedure implemented by Bocincova and Johnson (2019) as well as in numerous other EEG decoding studies (Foster et al., 2016; Fukuda et al., 2015; Sutterer et al., 2019; van Moorselaar et al., 2017). The baselined preprocessed EEG data were first band-pass filtered using a two-way least squares finite impulse response filter (EEGLAB function: "eegfilt.m"). Data were band-pass filtered in 1-Hz bands from 4 to 50 Hz (i.e., 4–5 Hz, 5–6 Hz, etc.) A Hilbert transform (Matlab Signal Processing Toolbox) was applied to the band-pass filtered data to obtain the complex analytic signal. Evoked power was calculated by first averaging the complex signal across trials within each training and test set (see Training and test data) and then, squaring the absolute value of the averaged signal. This evoked power signal reflects the activity that is phase locked to the stimulus onset because averaging before squaring the absolute value of the complex analytic signal cancels out signals that are out of phase. Total power was calculated by squaring the complex magnitude of the complex analytic signal and then, averaging across trials. This total power signal reflects ongoing activity regardless of its phase relationship to the onset of the chromatic disk. Note, that to avoid edge artifacts during our time window of interest (Cohen, 2014), a longer trial epoch of 1,000 ms before and 1,550 ms after stimulus onset was used for filtering and calculating instantaneous power. The extraneous time points were discarded before classification. To further reduce computation time for the classification analysis across time and frequency, the matrix of power values was down sampled to 1 sample every 20 ms. Power values (i.e., after filtering and applying the Hilbert transform) were down-sampled so that down-sampling did not affect the calculation of power.

2.10 | Pattern classification and partitioning data into training and test sets

Pattern classification was conducted on both ERP data and evoked power data. A naïve Bayes classifier implementation of a linear discriminant analysis ("classify" function in Matlab with "diaglinear" argument) was used to classify chromaticity, luminance, and joint chromaticity and luminance at each time point for analyses run on ERP data, and time–frequency point for analyses conducted on evoked and total power data.

For the pattern classification procedure, artifact free trials were partitioned into independent sets of training and test data for each observer. Across all analyses, trials were partitioned into three independent sets. The number of trials for each joint color and luminance value in each set were equated. Because of this constraint, a subset of trials was not assigned to any set. Thus, an iterative approach was used to make use of all available trials. For each iteration, trials were randomly partitioned into sets, as described, and pattern classification was performed on the resulting training and test data. Therefore, the trials that were not included in any set were different for each iteration. The resulting classification accuracies across iterations were averaged. This iterative approach reduced noise in the resulting classifier outputs by minimizing the influences of idiosyncrasies that were specific to any given assignment of trials to sets. For analyses focused on ERP data, 50 iterations were conducted. For trials focused on evoked power across a wide range of frequencies (which is a time-consuming procedure), 10 iterations were conducted.

Once trials were assigned to the three sets, averages for each stimulus feature of interest were calculated depending on the analysis (chromaticity, luminance, joint chromaticity and luminance) to obtain a matrix of ERPs or power values across all electrodes for each category set (electrodes × category sets, for each time point). A leave-one-out-cross-validation routine was used such that two of three sets served as training data and the remaining set served as the test data. For chromaticity, luminance, and joint chromaticity and luminance classification, the classifier routine was applied three times using each of the three matrices as the test set, and the remaining two as the training set. A slightly different approach was used for cross-validation of the cross-training analyses (see section "Cross-training analyses"). Different analyses required that the data for chromaticity and luminance categories be partitioned into training and test sets differently depending on the goal of the analysis. In the following subsections, we outline how data were partitioned into training and test sets for each analysis.

2.11 | Chromaticity and luminance classification

To allow for a more direct comparison in classifier activity across analyses (i.e., considering whether classification was "better" for chromaticity or luminance) the number of trials included in each set of the luminance and chromaticity only analyses were also equated. Specifically, the number of trials included in the chromaticity only analysis was reduced for Experiment 1, and the number of trials included in the luminance only analysis was reduced for Experiment 2. The number of trials included in each set for each condition was

equal for the combined analysis of Experiments 1 and 2, so neither needed to be reduced.

2.12 | Cross-training analyses

For analyses in which we assessed the similarity of patterns representing chromaticity at different luminance levels, data from each condition were again partitioned into three sets. For Experiment 1, the classifier was trained on two sets of chromaticity data that were an equal mix of two luminance levels and tested on the remaining set of chromaticity data for the held-out luminance level. The analysis was repeated so that the chromaticity at each luminance level served as the test set. The resulting accuracies were averaged across the held-out test sets. Similarly, for Experiment 2 and the combined analysis of Experiments 1 and 2, the classifier was trained on two sets of chromaticity data at one luminance level and the model was tested on one set of chromaticity data that consisted of only the untrained luminance level. The analysis was repeated so that chromaticity at each luminance level served as the test set, and the resulting accuracies were averaged across the held-out test sets.

For analyses in which we assessed the similarity of patterns representing luminance at different chromaticity levels, data from each condition were again partitioned into three sets. For Experiment 1 and the combined analysis of Experiments 1 and 2, the classifier was trained on two sets of luminance data at one chromaticity level and tested on one set of luminance data for the held-out chromaticity level. The analysis was repeated so that luminance at each chromaticity level served as the test set, and the resulting accuracies were averaged across held-out test sets. Similarly, for Experiment 2, the classifier was trained on two sets of luminance data at three chromaticity levels before the model was tested on the remaining set of luminance data that consisted of only the untrained chromaticity level. The analysis was repeated so that luminance at each chromaticity level served as the test set, and the resulting accuracies were averaged across held-out test sets. Note that the number of held-out conditions varied between the luminance and chromaticity cross-training analyses. Therefore, the number of trials included in each training and test set is not equated between the luminance and chromaticity cross-training analyses.

2.13 | Statistical analysis: Cluster-based permutation test

A cluster-based permutation test was used to identify when classifier accuracy was reliably above chance, while controlling for multiple comparisons (Cohen, 2014; Maris & Oostenveld, 2007). Clusters in which classifier accuracy was reliably above chance were identified for each analysis by performing a one-sided Wilcoxon signed-rank test (with the MATLAB "signrank" function) against chance level (i.e., one-sided sign rank against 0.5 for the chromaticity analysis, against 0.33 for the luminance analysis, and against .1667 for the joint chromaticity and luminance analyses in Experiment 1) at each time point in the ERP analyses (or at each time–frequency point in the time × frequency analysis). Next, clusters of contiguous time (ERP analysis) or time–frequency points (time × frequency analysis) that exceeded chance were identified. For each cluster, a test statistic was calculated by summing all rank (W) values in the cluster. A Monte Carlo randomization procedure was used to empirically approximate a null distribution for this test statistic. Specifically, the classification procedure was repeated 1,000 times for each subject but the category labels within each training and test set were randomized (see "Training and test data") so that the labels were random with respect to the observed response at each electrode. For each permutation, classifier accuracy was calculated across time or across time and frequencies to identify clusters as described above. For each permutation, the highest summed test statistic for any cluster was calculated resulting in a null distribution of 1,000 cluster test statistics. Finally, clusters that had test statistics larger than the 95th percentile of the null distribution were identified. Thus, the cluster test was a one-tailed test with an alpha level of 0.05 corrected for multiple comparisons.

3 | RESULTS

3.1 | Experiment 1

Differences in both chromaticity and luminance modulate VEPs, so an experimental challenge is to determine whether patterns of EEG activity track stimulus chromaticity, luminance, or both. In Experiment 1, we adopted a two-pronged approach to answer this question. First, we used HFP to equate the stimuli of different chromaticities in luminance. Second, we presented the chromatic stimuli at different luminance levels. This allowed a test of whether patterns of activity that allow decoding of chromaticity at one luminance level generalize to a substantially different luminance level.

3.1.1 | Behavioral performance

During the EEG task, observers were instructed to press a button if they detected that the stimulus was presented for longer than the usual duration of 100 msec. Average accuracy on the task was 95.7% (SD = 5.3), demonstrating that

observers successfully paid attention to the colored stimuli during the task.

3.1.2 | Patterns of EEG activity track differences in stimulus chromaticity

To test whether it is possible to decode the chromaticity of a presented stimulus from patterns of EEG activity, a linear discriminant classifier was used to decode the chromaticity ("red" or "green") of presented stimuli collapsing across luminance levels. If it is possible to decode chromaticity, then classifier accuracy should be higher than expected by chance (50%). Consistent with past work (Bocincova & Johnson, 2019), an initial step tested whether stimulus chromaticity could be decoded from patterns of evoked activity on the scalp across a range of different frequencies. Decoding of chromaticity was above chance for four clusters of evoked power ranging from 4 to 35 Hz within 500 ms of stimulus onset (Figure 2a; $p < .05$, cluster-based permutation test). Note, that above chance decoding prior to stimulus onset at low frequencies and toward the end of each trial likely reflects temporal smearing of EEG power due to time–frequency decomposition. Specifically, the power at a given time point for each frequency band in the time–frequency plots is calculated via a filter kernel that incorporates data from adjacent time points. The number of adjacent timepoints necessary for obtaining a power estimate at each time point in a particular frequency band is equal to the number of timepoints necessary to observe three cycles of the frequency band centered on the time point of interest. Thus, the number of adjacent time points that contribute to each pixel in Figure 2a decreases as frequency increases. For example, classification accuracy for a time point filtered at 4 Hz reflects activity from 375ms before and 375ms after the time point, classification accuracy for a time point filtered at 10 Hz reflects activity from 150 ms before and 150 ms after the time point, and classification accuracy for a time point filtered at 25 Hz reflects activity from 60 ms before and 60 ms after each timepoint. In contrast to decoding based on evoked power, decoding of chromaticity from patterns of total power was not above chance. Given that such a broad spectrum of frequencies support classification, and that chromatic information was carried by evoked power, we reasoned that this low frequency activity likely reflects the time–frequency make-up of the ERPs. Filtering necessarily reduces temporal precision, so classification was also run on the ERP data for each color category. Classifier accuracy was significantly higher than expected by chance for several clusters of time points from ~100 to 200 ms after stimulus presentation (Figure 2b; $p < .05$, cluster-based permutation test; $M = 70.1\%$, SD = 2.8% for significant time points), providing evidence that patterns of ERPs contain information about the chromaticity of presented stimuli. To determine if classification was driven by accurate identification of both "red" and "green" stimuli and not classifier bias, we plotted a confusion matrix (Figure 2c) of classifier output as a function of the presented stimulus. This confusion matrix is the average of all time points where classification of chromaticity was above chance (see Figure 2b). The confusion matrix (Figure 2c) revealed a clear peak in classification accuracy for both "red" and "green" stimuli confirming that both types of stimuli were accurately classified.

3.1.3 | Patterns of EEG activity track differences in stimulus luminance

To test whether it is possible to decode the presented luminance (6.5, 10.8, 17.4 cd/m^2) of a stimulus from patterns of EEG activity, a linear discriminant classifier was used to decode the luminance of presented stimuli collapsing across chromaticities. If it is possible to decode luminance, classifier accuracy should be higher than expected by chance (33%). Classifier accuracy was significantly higher than expected by chance across a somewhat narrower frequency range of evoked power (~15–25 Hz) than observed for chromaticity from ~100 to 300 ms after stimulus presentation ($p < .05$, cluster-based permutation test; Figure 2d). As with chromaticity, classifier accuracy did not exceed chance for any time–frequency clusters of total power. Classification of the ERPs revealed above chance decoding of luminance for several clusters of time points from ~200 to 650 ms after stimulus onset. (Figure 2e; $p < .05$, cluster-based permutation test; $M = 46.9\%$, SD = 1.8%, for significant time points). To determine whether each individual luminance level contributed to classifier performance, the average confusion matrix (Figure 2f) was calculated for all time points where classification accuracy was significantly above chance (Figure 2e). This confusion matrix (Figure 2f) revealed a clear peak in classification for each luminance level confirming that each luminance level could be accurately decoded. Together these results provide evidence that patterns of ERPs also contain information about stimulus luminance in addition to chromaticity and that these signals can be decoded while both types of information are simultaneously present in the stimulus.

3.1.4 | Patterns of EEG activity track joint differences in stimulus chromaticity and luminance

To test whether it is possible to decode the specific luminance and chromaticity pairing of a presented stimulus from patterns of ERPs, a linear discriminant classifier was trained

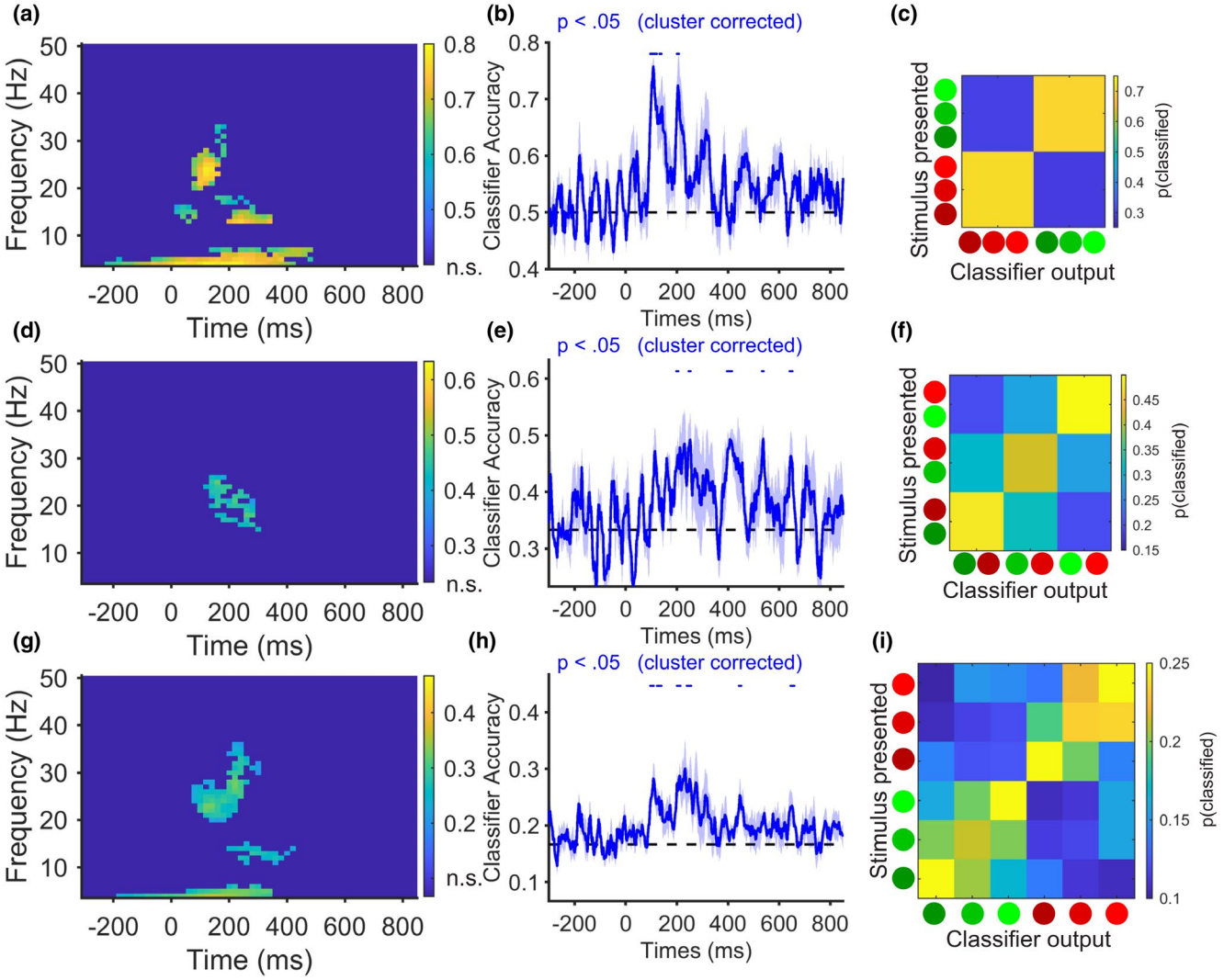

FIGURE 2 Classification of chromaticity, luminance, and joint chromaticity and luminance from topographic patterns of EEG activity for Experiment 1. (a) Decoding presented chromaticity from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (b) Decoding presented chromaticity from the topographic distribution of ERPs across time.** (c) Confusion matrix of chromaticity classifier choices averaged across significant ERP classification timepoints. (d) Decoding the presented luminance from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (e) Decoding presented luminance from the topographic distribution of ERPs across time.** (f) Confusion matrix of luminance classifier choices averaged across significant ERP classification timepoints. (g) Decoding the joint chromaticity and luminance of stimuli from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (h) Decoding the joint chromaticity and luminance of stimuli from the topographic distribution of ERPs across time.** (i) Confusion matrix of joint chromaticity and luminance classifier choices averaged across significant ERP classification timepoints. *Time–frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < .05$) were set to dark blue (the lowest value of the color scale). **Blue dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$). The shaded error bars reflect ± 1 SEM across observers

to decode each specific combination of chromaticity and luminance ("green" low luminance, "red" middle luminance, etc). If it is possible to decode the specific combination of chromaticity and luminance, classifier accuracy should be higher than expected by chance (16.7%). Classifier accuracy was significantly higher than expected by chance for three clusters of evoked power frequencies (4–35 Hz) within ~400 ms of stimulus onset (Figure 2g; $p < .05$, cluster-based permutation test). Once again, classifier accuracy did not exceed chance for any time–frequency clusters of total power. Because total power did not allow for decoding of color, luminance, or joint color and luminance, all subsequent analyses in Experiment 1 focused on patterns of evoked power and ERPs. Classification of patterns of ERPs revealed above chance decoding for several clusters of time points between 100 and 700 ms after stimulus onset (Figure 2h; $p < .05$, cluster-based permutation test; $M = 25.1\%$, SD = 2.2%, for significant time points), providing further evidence that patterns of ERPs contain information about the conjoined luminance and chromaticity of a stimulus.

Categorizing each chromaticity and luminance level provided an opportunity to compare the relative strengths of the underlying signals by analyzing the output of the classifier as a function of the presented stimulus (i.e., analyzing the confusion matrix). In other words, looking at how often the classifier made specific mistakes when it did not guess the correct category may provide useful information about the relative strength of the chromatic and luminance signals. If differences in chromaticity are more identifiable than luminance here, the expectation is that the classifier would more likely mistake "red low" for "red" at another luminance level, than for a different chromaticity at the same luminance like "green low." Alternatively, if luminance differences are more identifiable than chromatic differences, the classifier would be more likely to confuse stimuli of similar luminance levels than chromaticities (i.e., "red low" with "green low" rather than "red low" with "red medium"). Finally, it is possible that the strength of each signal is similar resulting in a similar proportion of confusions within chromaticity and luminance. Note that the relative strengths of chromaticity and luminance are specific for the particular chromaticities and luminances used here. While these chromaticities here were chosen to be easily discriminable from each other, and the luminance-level differences were far above threshold, the relative strength of chromaticity versus luminance here may not be assumed to generalize to other stimuli.

To test this question, an average confusion matrix was calculated for all time points where classification accuracy was significantly above chance (Figure 2h) and a binomial test was conducted on this group level output of the confusion matrix (Figure 2i). This allowed for a test of whether the classifier selected the correct chromaticity, but incorrect luminance, with higher probability than selecting any luminance level of the incorrect chromaticity, for each row of the confusion matrix (i.e., for a correct response of "medium red," did the classifier select both "low" and "high red" a higher proportion of the time, than "low," "medium," or "high green"). Assuming that all five misclassifications are equally likely, the probability of observing this pattern of results by chance for a given row is 12 of 120 possible orderings or $p = .10$. Misclassifications were more likely to be made within the same chromaticity than any luminance level for five of six rows of the confusion matrix. The binomial probability of observing this pattern of results for five or more rows by chance is $p < .001$ providing evidence that while both chromaticity and luminance information are present in patterns of ERPs, detection of chromatic differences was the dominant factor.

3.1.5 | Are chromatic and luminance patterns of EEG activity dissociable?

The confusion matrix results provide evidence that the signals for "red" and "green" chromaticities are strongly dissociable from each other. Classifier confusion values, however, are almost certainly dependent on the specific chromaticities, luminance values, and number of categories compared. Thus, comparing classifier confusions does not test whether the chromatic and luminance signals are dissociable. A better test of the ability to dissociate chromatic from luminance signals is to determine whether the chromatic signal observed generalizes across substantial changes in luminance. To test this hypothesis, a classifier was trained on stimulus chromaticity with data from two of three luminance levels, and tested on chromaticity at the *untrained* luminance level. The results showed classifier accuracy was significantly higher than expected by chance for multiple clusters of low frequency evoked activity across a range of frequencies spanning 4–50 Hz for up to 400 ms after stimulus onset (Figure 3a; $p < .05$, cluster-based permutation test). The same analysis on ERPs revealed above chance decoding of chromaticity for clusters of time points from ~100 to 800 ms after stimulus onset (Figure 3b; $p < .05$, cluster-based permutation test; $M = 70$, $SD = 7.3\%$ across significant time points). To determine if both "red" and "green" stimuli contributed to successful classification performance, the average confusion matrix (Figure 3c) was calculated for all time points where classification accuracy was significantly above chance (Figure 3b). This confusion matrix (Figure 3c) revealed a clear peak in classification accuracy for both "red" and "green" stimuli. Together these results provide clear evidence that patterns of ERPs contain chromatic information that can be measured despite substantial variations in luminance.

A related question is whether luminance changes generalize across changes in chromaticity. To test this question, the classifier was trained on luminance at one chromaticity level, and tested on luminance at the untrained chromaticity level. Classifier accuracy was significantly higher than expected by chance across two small clusters of evoked activity. One small significant cluster was observed from ~21 to 23 Hz within 200 ms of stimulus onset and another small significant cluster was observed at ~45–49 Hz roughly 600 ms after stimulus onset (Figure 3d; $p < .05$, cluster-based permutation test). Running the analysis on the ERPs showed above chance decoding of luminance for a single small cluster of time points roughly 400 ms after stimulus onset (Figure 3e; $p < .05$, cluster-based permutation test; $M = 53.2$, $SD = 4\%$ for significant time points). To determine whether each luminance level contributed to above chance classification performance, the average confusion matrix (Figure 3f) was calculated for all time points where classification accuracy was significantly above chance (Figure 3e). This confusion matrix (Figure 3f) revealed a clear peak in classification for each luminance level. While chromatic patterns of ERPs generalize across substantial changes in luminance, the evidence that luminance patterns of ERPs generalize across changes in chromaticity is relatively weaker for these stimuli.

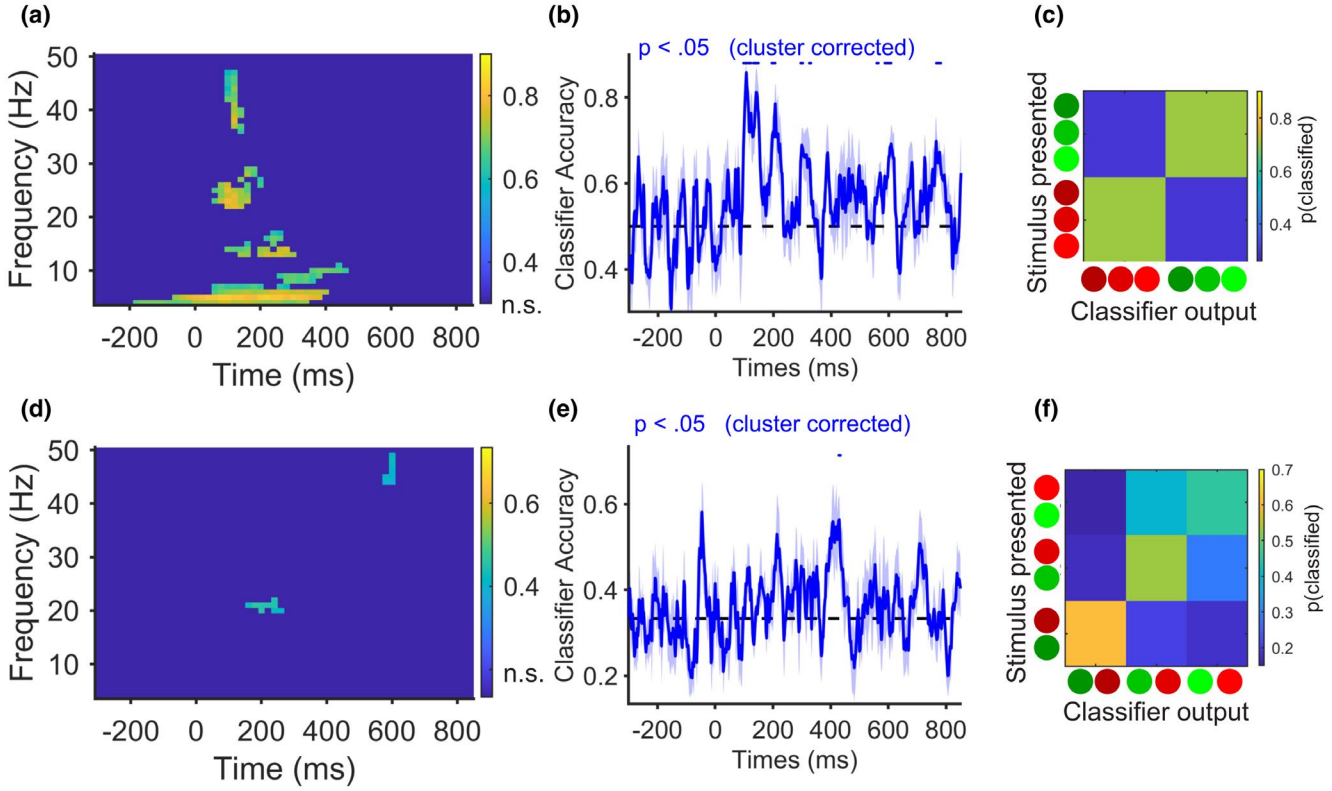

FIGURE 3 Chromatic patterns generalize across substantial differences in luminance. (a) Decoding of chromaticity from the topographic distribution of evoked power across a range of frequencies (4–50 Hz) when the classifier was trained and tested on stimuli of different luminance values.* (b) Decoding presented chromaticity from the topographic distribution of ERPs across time when the classifier was trained and tested on stimuli of different luminance levels.** (c) Confusion matrix of chromaticity classifier choices averaged across significant ERP classification timepoints when the classifier was trained and tested on stimuli of different luminance levels. (d) Decoding of luminance from the topographic distribution of evoked power across a range of frequencies (4–50 Hz) when the classifier was trained and tested on stimuli of different chromaticity values.* (e) Decoding presented luminance from the topographic distribution of ERPs across time when the classifier was trained and tested on stimuli of different chromaticity levels.** (f) Confusion matrix of luminance classifier choices averaged across significant ERP classification timepoints when the classifier was trained and tested on stimuli of different chromaticities. *Time–frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < 0.05$) were set to dark blue (the lowest value of the color scale). **Blue dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$). The shaded error bars reflect ±1 SEM across observers

3.2 | Experiment 2

Experiment 1 revealed that both chromaticity ("red" and "green") and luminance can be decoded from patterns of EEG activity. Furthermore, patterns of ERPs that differentiate "red" and "green" stimuli at one luminance level, generalize to "red" and "green" stimuli at a different luminance level. In addition to "red" and "green" chromatic stimuli, Experiment 2 included two additional chromaticities that fell along a continuum between "green" and "red" ("yellow" and "orange") at two different luminance levels. This allowed a replication of the results from Experiment 1 with a larger sample of participants, while also testing whether patterns of ERPs allow for chromatic classification of more similar chromaticities than used in Experiment 1.

3.2.1 | Behavioral performance

During the EEG task, observers were instructed to press a button if they detected that the stimulus was presented for longer than the usual duration (100 msec). Average accuracy on the task was 99.5% ($SD = .8\%$).

3.2.2 | Patterns of EEG activity track differences in stimulus chromaticity

To test whether it is possible to decode which of four chromaticities was presented from patterns of EEG activity, a linear discriminant classifier was used to decode the presented chromatic stimuli ("green," "yellow," "orange," or "red"), collapsing across the luminance levels. If it is possible to

decode chromaticity, then the classifier accuracy should be higher than expected by chance (25%). Above chance decoding of chromaticity was supported by evoked power from 4 to 45 Hz within 800 ms of stimulus onset (Figure 4a). In contrast to evoked power, decoding of chromaticity from total power was constrained to a narrower set of frequencies. Above chance decoding was observed for two clusters of total power ranging from 4 to 14 Hz within 600 ms of stimulus onset (Figure S1a; $p < .05$, cluster-based permutation test). When the classifier was trained on ERPs, we found that classifier accuracy was significantly higher than expected by chance from ~100 to 600 ms after stimulus presentation ($p < .05$, cluster-based permutation test; $M = 36.5\%$, SD = 4% for significant time points; Figure 4b).

One goal of Experiment 2 was to determine if chromaticities could be classified accurately using EEG even

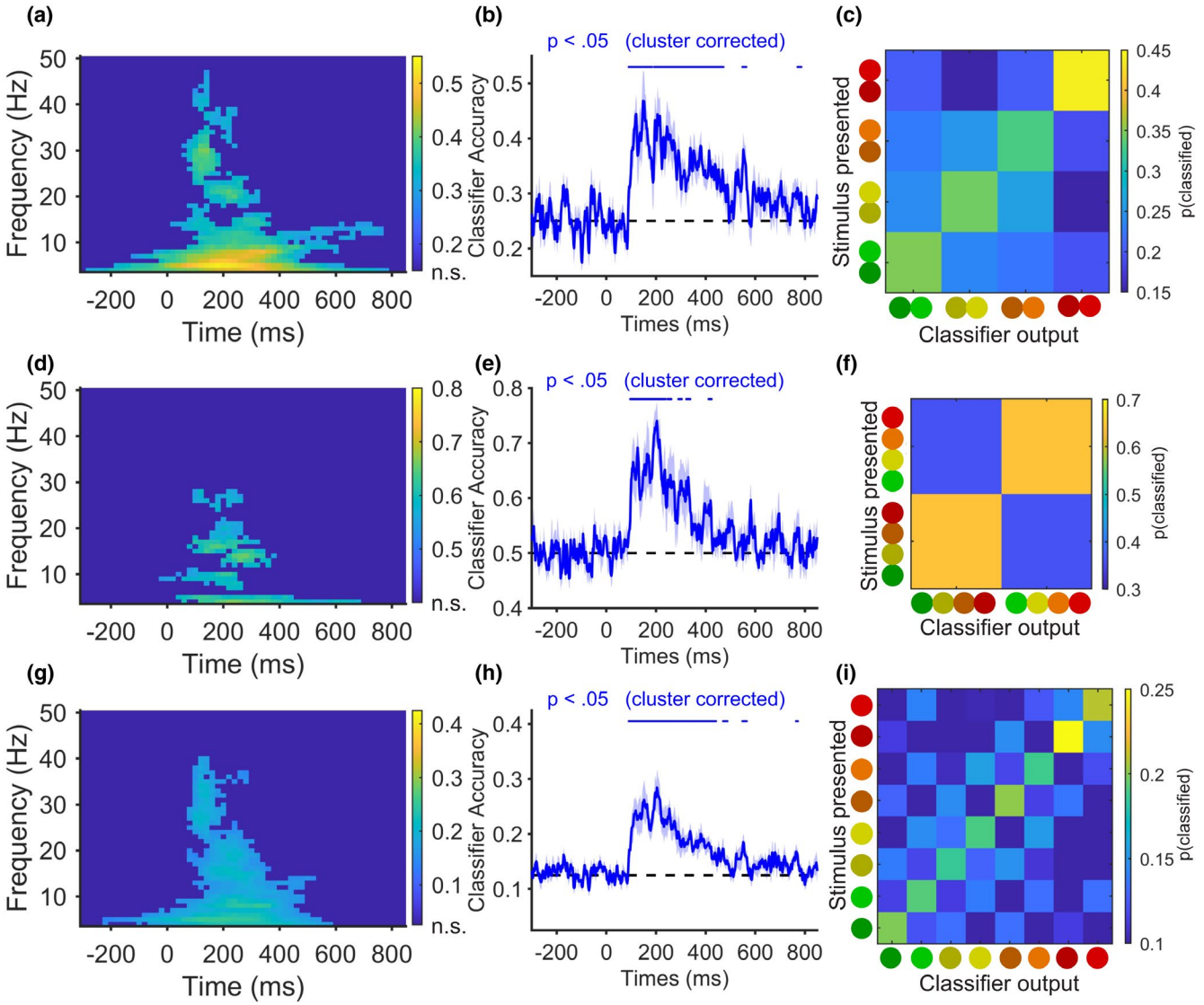

FIGURE 4 Classification of chromaticity, luminance, and joint chromaticity and luminance from topographic patterns of EEG activity for Experiment 2. (a) Decoding presented chromaticity from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (b) Decoding presented chromaticity from the topographic distribution of ERPs across time.** (c) Confusion matrix of chromaticity classifier choices averaged across significant ERP classification timepoints. (d) Decoding the presented luminance from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (e) Decoding presented luminance from the topographic distribution of ERPs across time.** (f) Confusion matrix of luminance classifier choices averaged across significant ERP classification timepoints. (g) Decoding the joint chromaticity and luminance of stimuli from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (h) Decoding the joint chromaticity and luminance of stimuli from the topographic distribution of ERPs across time.** (i) Confusion matrix of joint chromaticity and luminance classifier choices averaged across significant ERP classification timepoints. *Time–frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < .05$) were set to dark blue (the lowest value of the color scale). **Blue dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$). The shaded error bars reflect ±1 SEM across observers

if the chromaticities are not opponent colors (Shevell & Martin, 2017). If chromaticity classification can only be achieved with opponent colors, such as "red" and "green" as used in Experiment 1, we would expect that the classifier would be able to successfully categorize "red" versus "green" stimuli but be at chance for "red" versus "orange" stimuli. To determine whether classification performance was driven by each individual chromaticity or solely by "red" and "green" classification, the average confusion matrix (Figure 4c) was calculated for all time points where classification accuracy was significantly above chance (Figure 4b). This confusion matrix (Figure 4c) revealed a graded pattern of classification accuracy with a peak at each viewed chromaticity. Together these results provide evidence that patterns of ERPs contain information about the specific chromaticity of these stimuli and not only "red" and "green."

3.2.3 | Patterns of EEG activity track differences in stimulus luminance

To test whether it is possible to decode the luminance (6.5 or 10.8 cd/m^2) of a presented stimulus from the topography of ERPs, a linear discriminant classifier was used to decode the luminance of presented stimuli collapsing across chromaticities. If it is possible to decode luminance, classifier accuracy should be higher than expected by chance (50%). Once again, classifier accuracy was significantly higher than expected by chance though across a somewhat narrower range of evoked power than for chromaticity (~4–30 Hz) within 600 ms of stimulus presentation (Figure 4d; $p < .05$, cluster-based permutation test). In contrast to evoked power, decoding of stimulus luminance from total power did not exceed chance (Figure S1b). Classification of the ERPs revealed above chance decoding of luminance for clusters of time points from ~100 to 400 ms after stimulus presentation (Figure 4e; $p < 0.05$, cluster-based permutation test; $M = 64.2\%$, SD = 4% for significant time points), providing further evidence that patterns of ERPs contain information also about stimulus luminance. To determine whether both luminance levels contributed to classification performance, the average confusion matrix (Figure 4f) was calculated for all time points where classification accuracy was significantly above chance (Figure 4e). This confusion matrix (Figure 4f) revealed a clear peak in classification accuracy for both luminance levels.

3.2.4 | Patterns of EEG activity track joint differences in stimulus chromaticity and luminance

To test whether it is possible to decode the joint luminance and chromaticity of stimuli from the topography of ERPs, a linear discriminant classifier was used to decode each separate combination of presented stimuli. If it is possible to decode the specific combination of chromaticity and luminance (e.g., "green" low luminance, "orange" high luminance, etc.), classifier accuracy should be higher than expected by chance (12.5%).

As in Experiment 1, above chance decoding of the specific chromaticity and luminance of each stimulus was supported by a broad spectrum of evoked power (~4–40 Hz) within 600 ms of stimulus onset (Figure 4g; $p < .05$, cluster-based permutation test). Decoding of joint chromaticity and luminance from total power was constrained to a narrower set of frequencies and time points. Above chance decoding was observed for two clusters of total power ranging from 4 to 14 Hz within 600 ms of stimulus onset (Figure S1c; $p < .05$, cluster-based permutation test). Total power comprises both activity that is phase-locked to stimulus onset and non-phase locked activity (Cohen, 2014). The observations that classification of chromaticity and of joint chromaticity and luminance were constrained to a subset of times and frequencies tracked by evoked power suggest that classification based on total power also reflects activity phase-locked to the stimulus. Thus, as in Experiment 1, subsequent analyses in Experiment 2 focused on patterns of evoked power and ERPs. Decoding of ERPs was significantly higher than expected by chance from ~100 to 600 ms after stimulus presentation (Figure 4h; $p < .05$, cluster-based permutation test; $M = 20.3\%$, SD = 3.3% for significant time points), providing evidence that patterns of ERPs contain information about both the luminance and chromaticity of stimuli.

Once again, categorizing each chromaticity and luminance level provided the opportunity to compare the strength of the underlying luminance and chromaticity signal by analyzing the mistakes the classifier made as a function of each presented stimulus category. The average confusion matrix was calculated for all time points where classification accuracy was significantly above chance (Figure 4h). A binomial test was conducted to test whether the classifier selected the correct chromaticity, but incorrect luminance, with higher probability than selecting any luminance level of the incorrect chromaticity, for each row of the confusion matrix (Figure 4i). Assuming that all seven misclassifications are equally likely, the probability of observing this pattern of results by chance for a given row is 720 of 5,040 possible orderings or $p = .14$. Misclassifications were more likely to be made within the same chromaticity than any luminance level for one of eight rows of the confusion matrix. The binomial probability of observing this pattern of results for one or more rows by chance was $p = .71$ so it fails to show that chromatic differences were dominant over luminance differences in Experiment 2. A second binomial test was conducted to assess whether the classifier selected the correct luminance, but incorrect chromaticity, with higher probability than selecting

any chromaticity of the incorrect luminance. Assuming that all seven misclassifications are equally likely, the probability of observing this pattern of results by chance for a given row is 144 of 5,040 possible orderings or $p = .0286$. Misclassifications were more likely to be made within the same luminance than any chromaticity level for three of eight rows of the confusion matrix. The binomial probability of observing this pattern of results for three or more rows by chance was $p = .0012$ providing evidence that luminance differences were dominant over chromaticity differences in Experiment 2. This finding is not necessarily unexpected, however, because as noted above the relative strengths of chromaticity and luminance depend on the particular chromaticities and luminances used in a given experiment, and Experiment 2 had stimuli with more subtle variations in chromaticity.

3.2.5 | Are chromatic and luminance patterns of EEG activity dissociable?

While chromaticity was not found to be a stronger signal than luminance in Experiment 2, an open question is whether the chromaticity signal generalizes across changes in luminance. As in Experiment 1, the classifier was trained on stimulus chromaticity at one luminance level, and tested at the untrained luminance level. Classifier accuracy was significantly higher than expected by chance across four clusters of evoked activity. Three significant clusters were observed from ~4 to 18 Hz within 600 ms of stimulus onset, and another significant cluster was observed from ~25 to 35 Hz and from ~50 to 200 ms after stimulus onset (Figure 5a; $p < .05$, cluster-based permutation test). Running the analysis on the ERPs showed above chance decoding of chromaticity for clusters of time points from ~150 to 600 ms after stimulus onset (Figure 5b; $p < .05$, cluster-based permutation test; $M = 37.3\%$, $SD = 3.4$ for significant time points). To determine whether classification performance of this generalizable chromaticity signal reflects decoding of each individual chromaticity or was driven by only "red" and "green" classification, the average confusion matrix (Figure 5c) was calculated for all time points where classification accuracy was significantly above chance (Figure 5b). This revealed a graded pattern of classification accuracy with a peak at each observed chromaticity (Figure 5c). Together these results provide evidence that patterns of ERPs contain information about the specific chromaticities of the stimuli that is dissociable from the luminance of the stimulus.

Another question is whether luminance changes generalize across changes in chromaticity. The classifier was trained on stimulus luminance at three chromaticity levels, and tested on luminance at the untrained chromaticity level. Classifier accuracy was significantly higher than expected by chance across three clusters of evoked activity. One significant cluster was observed from ~7 to 20 Hz within 600 ms of stimulus onset, another small significant cluster was observed at 4–5 Hz from ~200 to 600 ms after stimulus onset, and a third significant cluster was observed from ~25 to 35 Hz for ~ 50–150 ms after stimulus onset ($p < .05$, cluster-based permutation test; Figure 5d). Running the analysis on the ERPs showed above chance decoding of luminance for clusters of time points from ~100 to 500 ms after stimulus onset (Figure 5e; $p < .05$, cluster-based permutation test; $M = 66\%$, $SD = 3.5\%$ for significant time points). To determine whether both luminance levels contributed to above chance classification, the average confusion matrix (Figure 5f) was calculated for all time points where classification accuracy was significantly above chance (Figure 5e). This confusion matrix (Figure 5f) revealed a clear peak in classification accuracy for both luminance levels. Together, these results provide evidence that patterns of ERPs contain dissociable information about the specific chromaticity and luminance of each stimulus.

3.2.6 | Are differences in decoding timing and strength between experiments attributable to differences in task design or sample size?

When comparing results across experiments, several differences stand out. In particular, luminance decoding in Experiment 1 appeared weaker and reached significance later (~200 ms after stimulus onset; Figure 2e) than in Experiment 2 (~100 ms after stimulus onset; Figure 4e). This difference was also evident in the observation that chromaticity was the dominant signal in Experiment 1 (Figure 2i) while luminance was the dominant signal in Experiment 2 (Figure 4i). Additionally, luminance patterns showed weak generalization across changes in chromaticity in Experiment 1 (Figure 3d,e) and robust generalization in Experiment 2 (Figure 5d,e). Thus, an open question is the extent to which these inconsistencies between experiments reflect differences in task design (i.e., the number of chromaticies and luminance levels compared) and how much these differences reflect differences in experimental power (Experiment 1 had five observers while Experiment 2 had nine observers). In order to gain some insight into this question, we combined data across experiments by focusing only on stimulus values that were included in both experiments ("red" and "green" stimuli with low and medium luminance) and re-ran the core multivariate analyses. Time–frequency analysis of the combined data set appeared more similar to the results observed in Experiment 2, with a broad range of frequencies representing the chromaticity (Figure 6a), luminance (Figure 6d), and joint chromaticity

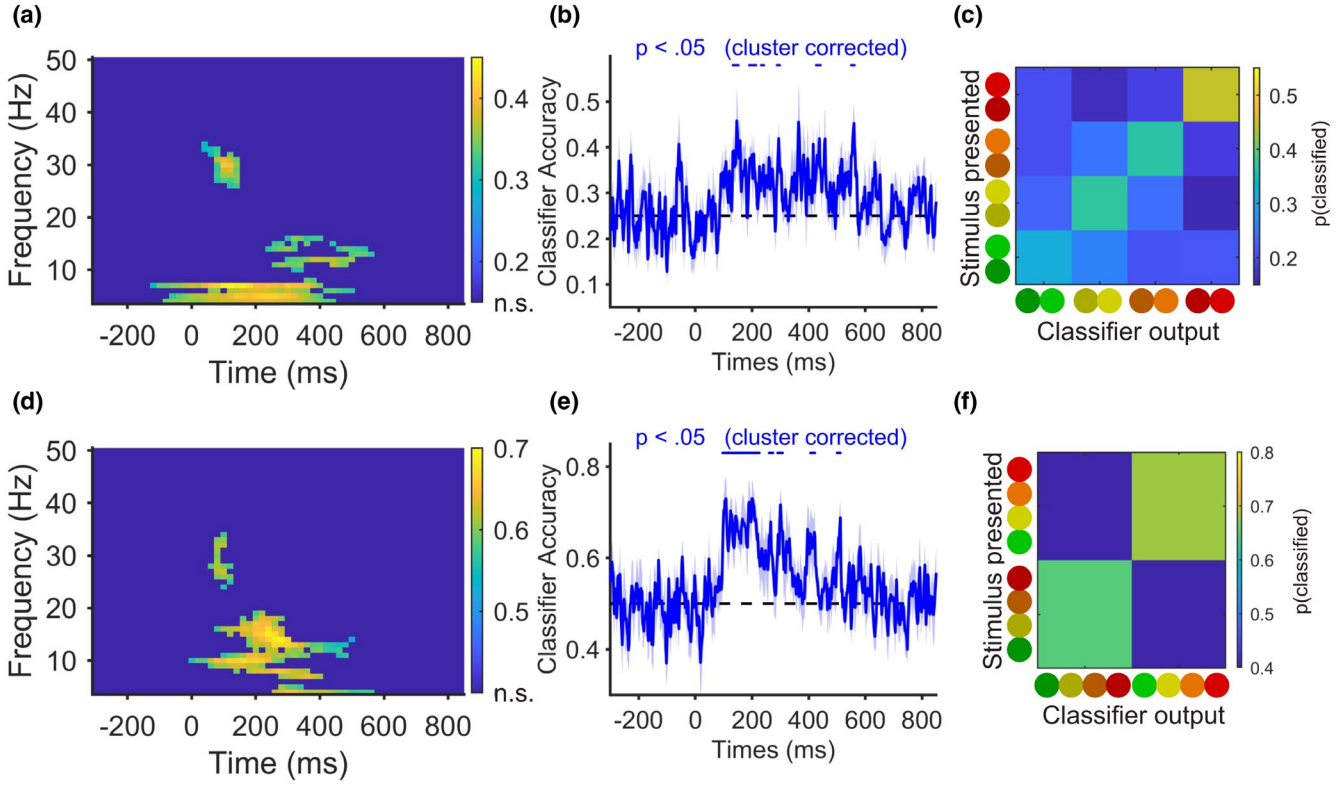

FIGURE 5 Chromatic patterns generalize across differences in luminance and vice versa in Experiment 2. (a) Decoding of chromaticity from the topographic distribution of evoked power across a range of frequencies (4–50 Hz) when the classifier was trained and tested on stimuli of different luminance values.* (b) Decoding presented chromaticity from the topographic distribution of ERPs across time when the classifier was trained and tested on stimuli of different luminance levels.** (c) Confusion matrix of chromaticity classifier choices averaged across significant ERP classification timepoints when the classifier was trained and tested on stimuli of different luminance levels. (d) Decoding of luminance from the topographic distribution of evoked power across a range of frequencies (4–50 Hz) when the classifier was trained and tested on stimuli of different chromaticity values.* (e) Decoding presented luminance from the topographic distribution of ERPs across time when the classifier was trained and tested on stimuli of different chromaticity levels.** (f) Confusion matrix of luminance classifier choices averaged across significant ERP classification timepoints when the classifier was trained and tested on stimuli of different chromaticities. *Time–frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < .05$) were set to dark blue (the lowest value of the color scale). **Blue dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$). The shaded error bars reflect ±1 SEM across observers

and luminance (Figure 6g) of presented stimuli. As before, however, a slightly more limited range of frequencies represented stimulus luminance. Analysis of total power also revealed a pattern of results similar to Experiment 2 for chromaticity and joint chromaticity and luminance, where fewer clusters of above chance decoding were observed for total than evoked power (Figures S1d,f). Several above chance clusters were observed for luminance as well (Figure S1e), suggesting that the absence of above chance classification of luminance from total power in Experiments 1 and 2 can be attributed to a combination of sample size and weaker decoding for total than evoked power. As in Experiment 2, decoding based on the pattern of ERPs revealed above chance classification of chromaticity (Figure 6b), luminance (Figure 6e), and joint chromaticity and luminance (Figure 6h) approximately 100 ms after stimulus onset. Classifier confusions revealed a clear peak in pattern classification for both "red" and "green" stimuli (Figure 6c) and low and medium luminance stimuli (Figure 6f) showing that both categories of chromaticity and luminance contributed to above chance decoding. This pattern of results confirms that it is possible to obtain robust decoding of "red" and "green" stimuli in the absence of intermediate chromaticities, and suggests that differences in luminance decoding strength and timing between Experiment 1 and Experiment 2 were likely due to the smaller sample size of Experiment 1. It is also possible that the inclusion of a third luminance level in Experiment 1 further weakened luminance decoding performance.

To determine whether differences between experiments in the extent that the chromatic signal (Experiment 1) or luminance signal (Experiment 2) dominated joint chromatic and luminance classifier performance could be attributed to differences in stimulus properties or experimental power, an average confusion matrix was calculated for all time points

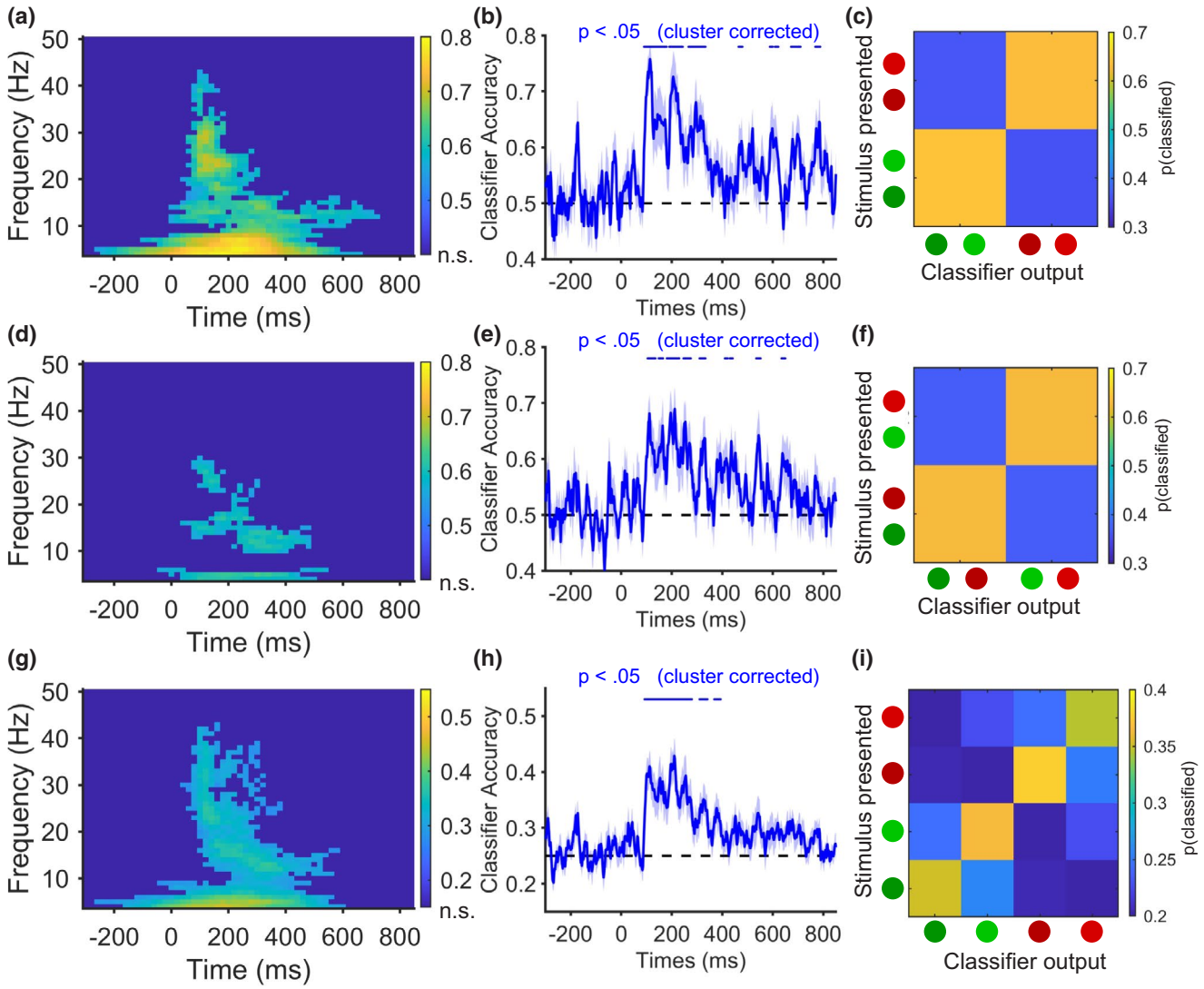

FIGURE 6 Classification of chromaticity, luminance, and joint chromaticity and luminance from topographic patterns of EEG activity for Experiments 1 and 2. (a) Decoding presented chromaticity from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (b) Decoding presented chromaticity from the topographic distribution of ERPs across time.** (c) Confusion matrix of chromaticity classifier choices averaged across significant ERP classification timepoints. (d) Decoding the presented luminance from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (e) Decoding presented luminance from the topographic distribution of ERPs across time.** (f) Confusion matrix of luminance classifier choices averaged across significant ERP classification timepoints. (g) Decoding the joint chromaticity and luminance of stimuli from the topographic distribution of evoked power across a range of frequencies (4–50 Hz).* (h) Decoding the joint chromaticity and luminance of stimuli from the topographic distribution of ERPs across time.** (i) Confusion matrix of joint chromaticity and luminance classifier choices averaged across significant ERP classification timepoints. *Time–frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < .05$) were set to dark blue (the lowest value of the color scale). **Blue dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$). The shaded error bars reflect ±1 SEM across observers

where joint chromaticity and luminance classification accuracy was significantly above chance (Figure 6i). The pattern of confusions observed (Figure 6i) was similar to that observed in Experiment 1 in which chromatic confusions were more likely than luminance confusions. A binomial test was conducted to test whether the classifier selected the correct chromaticity, but incorrect luminance, with higher probability than selecting any luminance level of the incorrect chromaticity, for each row of the confusion matrix (Figure 6i). Assuming that all three misclassifications are equally likely, the probability of observing this pattern of results by chance for a given row is two of six possible orderings or $p = .33$. Misclassifications were more likely to be made within the same chromaticity than any luminance level for four of four rows of the confusion

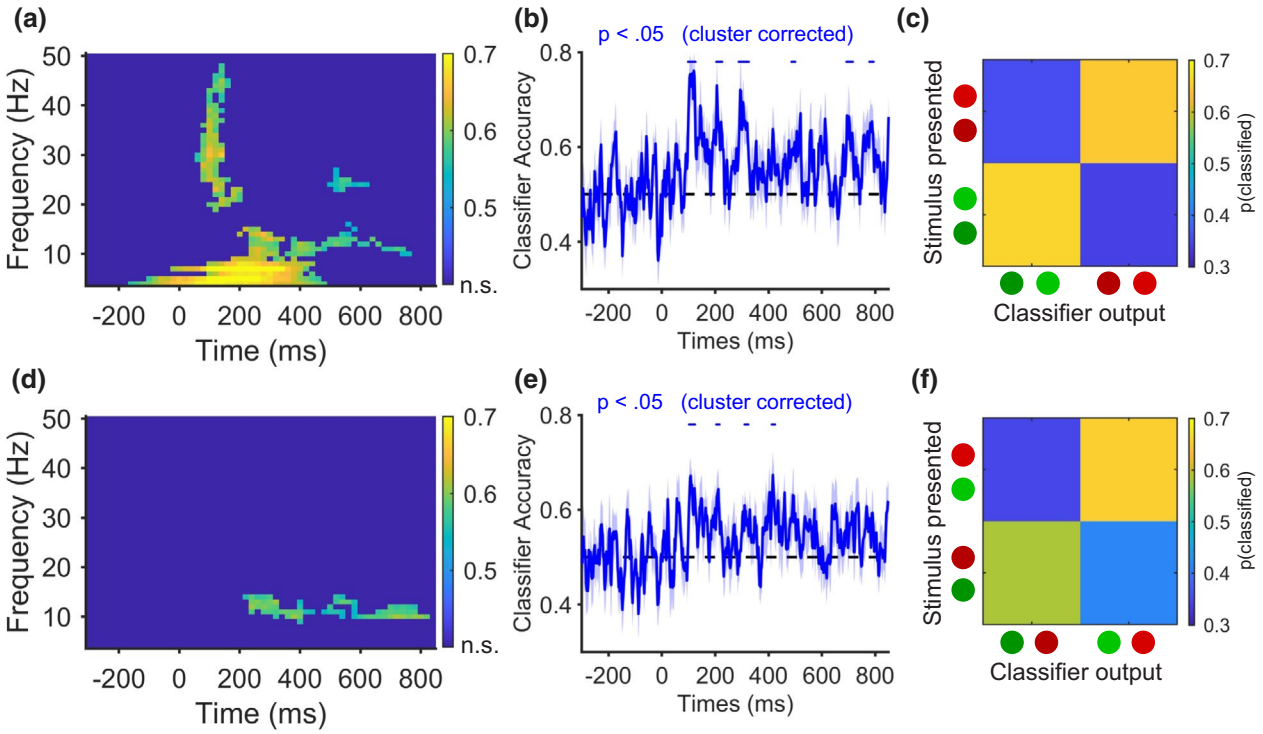

FIGURE 7 Chromatic patterns generalize across differences in luminance and vice versa in Experiments 1 and 2. (a) Decoding of chromaticity from the topographic distribution of evoked power across a range of frequencies (4–50 Hz) when the classifier was trained and tested on stimuli of different luminance values.* (b) Decoding presented chromaticity from the topographic distribution of ERPs across time when the classifier was trained and tested on stimuli of different luminance levels.** (c) Confusion matrix of chromaticity classifier choices averaged across significant ERP classification timepoints when the classifier was trained and tested on stimuli of different luminance levels. (d) Decoding of luminance from the topographic distribution of evoked power across a range of frequencies (4–50 Hz) when the classifier was trained and tested on stimuli of different chromaticity values.* (e) Decoding presented luminance from the topographic distribution of ERPs across time when the classifier was trained and tested on stimuli of different chromaticity levels.** (f) Confusion matrix of luminance classifier choices averaged across significant ERP classification timepoints when the classifier was trained and tested on stimuli of different chromaticities. *Time–frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < .05$) were set to dark blue (the lowest value of the color scale). **Blue dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$). The shaded error bars reflect ±1 SEM across observers

matrix. The binomial probability of observing this pattern of results for four rows by chance was $p = .012$ revealing that chromatic differences were dominant over luminance differences in the combined data set. The pattern of classifier confusions observed across analyses suggests that the observed differences in chromatic and luminance signal dominance between Experiments 1 and 2 are due to the chromaticities included in the comparison rather than differences in sample size or the number of luminance levels. Thus, this combined analysis provides further evidence that patterns of ERPs are sensitive to both stimulus chromaticity and luminance, and the comparative strength of these two signals is specific to the particular chromaticities and luminances used and should not be assumed to generalize to other stimuli.

To determine whether the weak generalization of luminance classification across chromaticities observed in Experiment 1 reflects low power or that the luminance signal does not generalize for "red" and "green" stimuli, the generalization analyses were also re-run on the combined data. Time–frequency analysis revealed that a broad range of frequencies (Figure 7a) tracked stimulus chromaticity across variation in luminance, while a narrower cluster of frequencies tracked stimulus luminance across variation in chromaticity (Figure 7d). Decoding from the topographic distribution of ERPs revealed above chance classification of chromaticity across substantial changes in luminance (Figure 7b) and above chance classification of luminance (Figure 7d) across different chromaticities. Confusion matrices revealed a clear peak in classification accuracy for both chromaticity levels (Figure 7c) and both luminance levels (Figure 7f). Together, these results suggest that the weak generalization of luminance decoding observed in Experiment 1 (Figure 3e,f) can be attributed to the

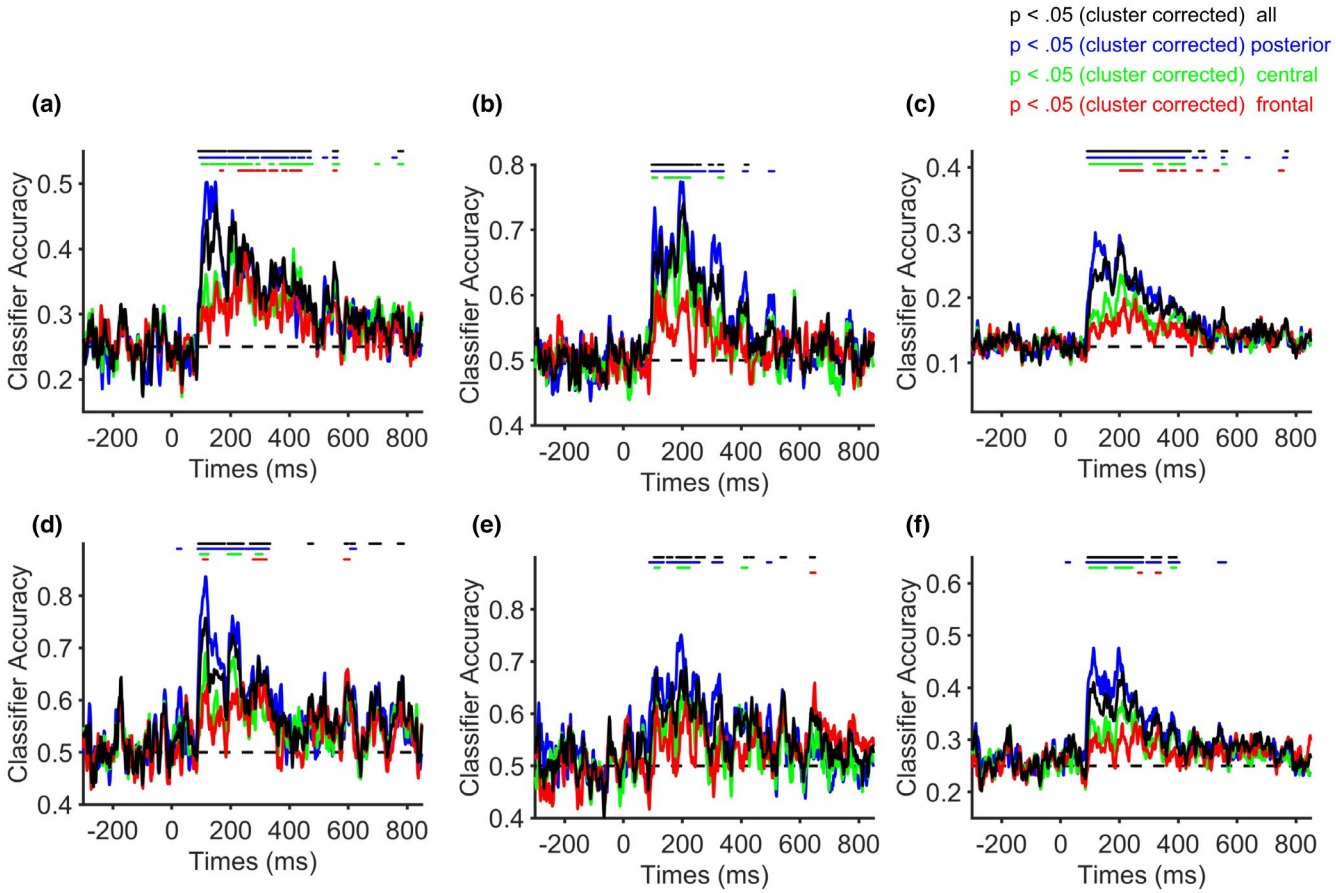

FIGURE 8 Comparing decoding accuracy at posterior, central, frontal, and all electrodes. (a) Decoding chromaticity from the topographic distribution of VEPs across time for posterior, central, frontal, and all electrodes for Experiment 2. (b) Decoding luminance from the topographic distribution of VEPs across time for posterior, central, frontal, and all electrodes for Experiment 2. (c) Decoding joint chromaticity and luminance from the topographic distribution of VEPs across time for posterior, central, frontal, and all electrodes for Experiment 2. (d) Decoding chromaticity from the topographic distribution of VEPs across time for posterior, central, frontal, and all electrodes for combined data from Experiments 1 and 2. (e) Decoding luminance from the topographic distribution of VEPs across time for posterior, central, frontal, and all electrodes for combined data from Experiments 1 and 2. (f) Decoding joint chromaticity and luminance from the topographic distribution of VEPs across time for posterior, central, frontal, and all electrodes for combined data from Experiments 1 and 2. Colored dots mark timepoints where classifier accuracy was significantly above chance as determined by a cluster corrected permutation test ($p < .05$)

modest sample size of Experiment 1 (n = 5) and possibly the inclusion of a third luminance level, rather than the absence of a generalizable luminance signal for "red" and "green" stimuli.

3.2.7 | Which electrodes contribute to decoding?

The ERP signal in both Experiment 2 and the combined analysis of Experiments 1 and 2 supported above chance decoding of chromaticity and luminance for up to 800 ms after stimulus onset, which is longer than typically observed for traditional VEPs. This raises the question of whether posterior electrodes that are typically associated with VEPs support decoding during that entire period, or if instead more anterior electrodes support decoding at later time points. The average voltage topography from 100 to 500 ms in Experiment 2 and the combined analysis of Experiments 1 and 2, revealed subtle differences in voltage across the scalp for stimuli of different chromaticies and luminances (Figure S2). These differences were also apparent in the VEPs of posterior, central, and frontal midline electrodes (Figure S3). To determine if and when posterior (Oz, O1, O2, PO3, PO4, PO7, and PO8), central (P3, P4, P7, P8, Pz, CP1, CP2, CP5, CP6, C3, C4, and Cz), and frontal (FC1, FC2, FC5, FC6, F3, F4, F7, F8, Fz, Fp1, and Fp2) electrodes contributed to decoding of chromaticity and luminance, linear discriminant classification was conducted on the patterns of ERPs from these three subsets of electrodes. Electrode groups were selected to ensure roughly equal representation from posterior to anterior across the scalp while including similar numbers of electrodes in each group. Linear discriminant classification revealed above chance classification of chromaticity

and joint chromaticity and luminance at all three sets of electrodes in both Experiment 2 and the combined data set (Figure 8a,c,d,f), providing evidence that a wide range of electrodes track stimulus chromaticity and joint chromaticity and luminance. However, above chance decoding of luminance from frontal electrodes was observed for only a small cluster of time points in the combined data set and was not observed in Experiment 2 (Figure 8b,e). In line with the substantial literature showing that posterior electrodes are most responsive to stimulus chromaticity and luminance, decoding accuracy was highest at posterior electrodes, followed by central electrodes, and was lowest at frontal electrodes. Interestingly, patterns of EEG activity at posterior electrodes tracked stimulus and chromaticity and luminance through a similar time window observed for classification using all electrodes on the scalp (Figure 8), suggesting that anterior electrodes were not the sole driver of sustained decoding. Finally, decoding accuracy for posterior electrodes was most similar to the accuracy observed when all electrodes were included, showing that even though classification accuracy was lower for more anterior electrodes, including them in the analysis did not greatly impair classification.

4 | DISCUSSION

Color vision has been studied with EEG for more than 50 years, and a long-standing question is whether it is possible to decode the chromaticity of a stimulus from patterns of EEG activity on the scalp. Past work has shown that discriminable VEPs are found with stimuli that vary in chromaticity or luminance (Murray et al., 1987; Paulus et al., 1984, 1986; Rabin et al., 1994; Skiba et al., 2014). Additionally, VEP waveforms have been proposed to differentiate between color categories such as whether or not a chromatic stimulus is a unique hue (Forder et al., 2017). However, whether differences observed across categories reflect a purely perceptual response to the stimuli remains a topic of ongoing debate (Siuda-Krzywicka et al., 2019), and these findings do not reveal whether EEG activity contains the information needed to discriminate specific chromaticity values (i.e., "red" or "green") in the absence of confounds with luminance.

We addressed this question using a multivariate pattern classification approach in which a classifier was trained to identify specific chromaticities and then, tested on a held-out data set (Brouwer & Heeger, 2009). It is well known that VEPs can respond to differences in either stimulus chromaticity or luminance (Kulikowski et al., 1996; Skiba et al., 2014), and previous EEG pattern classification work (Bocincova & Johnson, 2019; Rasheed & Marini, 2015; Sandhaeger et al., 2019) did not precisely match chromatic stimuli in luminance or test whether classification of chromaticity is maintained despite differences in luminance of the chromatic stimuli. Thus, our goal was to clarify whether EEG activity contains information about stimulus chromaticity per se.

We conducted two experiments in which observers monitored centrally presented chromatic disks that varied both in chromaticity and luminance while EEG was recorded. Pattern classification allowed for successful decoding of both stimulus chromaticity and luminance. The earliest time points allowing for decoding of chromaticity and luminance in our EEG analyses were between 90 and 100 ms (see Figure 6b,d,h). This observation is in line with past observations that early ERP components such as the n87 are sensitive to both stimulus chromaticity and luminance (Paulus et al., 1986), and matches up with the timing with which color information could be decoded from invasive recordings in primates and previous EEG and MEG from humans (Sandhaeger et al., 2019). Critically, we found that the chromaticity of stimuli could be decoded when the classifier was trained on the chromaticities presented at one luminance level and tested at a different luminance level. Thus, the topography of EEG activity can be used to track the precise chromaticity of a stimulus, and this chromatic signal generalizes across substantial changes in luminance. This finding is in line with a similar observation of a chromatic signal that generalizes across luminance levels with a single magneto encephalography (MEG) participant (Sandhaeger et al., 2019) suggesting that MEG and EEG are both sensitive to chromatic information.

Both chromaticity and luminance could be decoded from a relatively broad spectrum of evoked EEG power, which reflects activity that is phase-locked to the stimulus presentation. One question is whether chromaticity and luminance could be decoded from ongoing oscillations that are not phase locked to stimulus presentation in addition to the robust decoding we observed. We found no evidence for decoding in total power in any frequency from 4 to 50 Hz in Experiment 1, for chromaticity, luminance, or joint chromaticity and luminance. We did find evidence for decoding in several clusters of total power for chromaticity, luminance (combined data only), and joint chromaticity and luminance in Experiment 2 and the combined data from Experiments 1 and 2 (Figure S1). This total power signal reflects both ongoing oscillations and oscillations phase locked to the stimulus (Cohen, 2014). Thus, the observation that total power classification was less robust than evoked power suggests that chromaticity and luminance decoding are primarily supported by evoked power. A related question is whether or not chromatic and luminance signals vary systematically in the frequencies of evoked activity that support each of them. When the classifier was trained and tested on separate chromaticies, luminance activity appeared to be marked by frequencies tending toward the alpha–beta range (~7–20 Hz; Figures 5d and 7d) and chromatic activity tended to appear across a wider range

of high and low frequencies (Figures 5a and 7a). However, this analysis required the classifier to tolerate substantial variability in one signal to decode the other and both chromatic and luminance information could be decoded from a wide range of frequencies when the classifier did not have to generalize across chromaticity or luminance differences (Figures 4a,d and 6a,d). Thus, more work is needed to determine whether or not there are systematic differences in the frequencies that track stimulus chromaticity and luminance.

Another open question is where in the brain these chromatic EEG signals originate. Functional neuroimaging work finds that color information can be decoded throughout the cortical hierarchy as early as V1 and as far upstream as frontal cortex (Bird et al., 2014; Brouwer & Heeger, 2009, 2013; Kim et al., 2020; Siuda-Krzywicka et al., 2019). Recent work comparing color-specific patterns of activity from invasive cortical recordings in primates with patterns of activity from MEG recordings in humans suggests that MEG decoding of color information in humans is present throughout the cortical hierarchy but is strongest in early visual areas (Sandhaeger et al., 2019). One way to better understand which brain regions contribute to this EEG signal is to test the level of color organization that these signals reflect. For example, Brouwer and Heeger (2009) demonstrated that color information is represented differently in patterns of bold activity in V1 than in V4 with patterns in V1 reflecting low level processing of an opponent color space and patterns in V4 reflecting perceptual color space. The combined luminance matching and variation method used here could be extended to include more chromaticies in future work to allow similar measurement of chromatic organization.

The goal of the present work was to determine if EEG activity can be used to differentiate various chromaticities uncontaminated by possible differences in stimulus luminance. Whether or not the patterns of EEG activity that facilitate decoding of chromaticity are driven more by saturation or hue (percepts observers experienced) is an intriguing question, but beyond the scope of the current study. The chromaticities used here can be re-expressed in the CIELUV (u′, v′) coordinate system ("green" = 0.13,0.55, "yellow" = 0.20,0.54, "orange" = 0.33,0.52, "red" = 0.43,0.51), which approximates a uniform color scale in which differences between points in the color space are intended to correspond to visual differences between the colors seen. While these values may be useful for speculating about percepts observers experienced, the measurements here are neural responses to chromatic lights stimulating the retina and cannot be used to infer a role for the color percepts. Future work using an approach such as binocular switch rivalry (Kim et al., 2020), which can cause changes in color without altering the physical stimulus presented to the retina could be useful for answering this question.

Here we observed above chance decoding of chromatic and luminance information for up to 800 ms after stimulus onset, which is later than VEPs are typically observed. Recent work has shown that patterns of EEG activity track the online maintenance of object features in memory. For instance, the topography of total alpha-band (8–12 Hz) power can be used to track the active maintenance of remembered locations (Foster et al., 2016; Sutterer et al., 2019) while patterns of ERP activity track the maintenance of remembered orientations in working memory (Bae & Luck, 2017, 2019; Wolff et al., 2017). Could this relatively late decoding of chromatic and luminance information reflect maintenance of information in working memory even though there was no task demand for observers to remember this information? Bocincova and Johnson (2019) recently tested this question by assessing whether the color of an oriented grating could be tracked with EEG over a longer memory delay (1.8 s) while observers attempted to remember the color. However, similar to the joint chromaticity and luminance decoding in the present work, the color of the stimulus could be tracked for ~600 ms after stimulus onset but not over the remainder of the memory delay. These results suggest that the relatively late activity observed in the present experiment does not reflect memory maintenance, but instead reflects a fairly long-lasting response to the stimulus. Another potential explanation for the prolonged decoding observed in the present work is that stimuli were presented for a short duration (100 ms) with an abrupt onset and offset. Thus, the sustained chromatic and luminance decoding we observed may reflect a response to both the onset and offset of the stimulus in addition to decoding the properties of the disk when it was on the screen. The extent to which different cognitive operations (e.g., memory and attention) and stimulus properties (e.g., whether the stimulus is presented transiently or for a prolonged duration) affect the strength and duration of decoding are interesting questions for future research.

5 | CONCLUSIONS

A critical open question in the field of vision research is whether the chromaticity of a stimulus can be decoded from scalp-recorded EEG activity even when differences in luminance do not inform the classification. The approach here employed a two-step strategy of carefully equating chromatic stimuli in luminance and then, systematically varying both the chromaticity and luminance of stimuli. Then, a multivariate pattern classification analysis tested for patterns of activity that generalized across differences in luminance. This approach revealed that patterns of EEG activity can be used to decode the specific stimulus chromaticity, independent of variations in luminance.

ACKNOWLEDGMENTS

This work was supported in part by a shared equipment grant from the Neuroscience Institute at the University of Chicago.

AUTHOR CONTRIBUTION

David Sutterer: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Validation; Visualization; Writing-original draft; Writing-review & editing. **andrew coia:** Conceptualization; Data curation; Investigation; Methodology; Project administration; Software; Writing-original draft; Writing-review & editing. **Vincent Sun:** Conceptualization; Data curation; Investigation; Methodology; Software; Writing-review & editing. **Steven Shevell:** Conceptualization; Investigation; Methodology; Project administration; Resources; Supervision; Writing-review & editing. **Edward Awh:** Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing-review & editing.

ORCID

David W. Sutterer https://orcid.org/0000-0001-7977-6861
Andrew J. Coia https://orcid.org/0000-0002-1161-189X

REFERENCES

Bae, G.-Y., & Luck, S. J. (2017). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *The Journal of Neuroscience*, *38*(2), 409–422. https://doi.org/10.1523/jneurosci.2860-17.2017

Bae, G.-Y., & Luck, S. J. (2019). Reactivation of previous experiences in a working memory task. *Psychological Science*, *30*(4), 587–595. https://doi.org/10.1177/0956797619830398

Bird, C. M., Berens, S. C., Horner, A. J., & Franklin, A. (2014). Categorical encoding of color in the brain. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(12), 4590–4595. https://doi.org/10.1073/pnas.1315275111

Bocincova, A., & Johnson, J. S. (2019). The time course of encoding and maintenance of task-relevant versus irrelevant object features in working memory. *Cortex*, *111*, 196–209. https://doi.org/10.1016/j.cortex.2018.10.013

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *The Journal of Neuroscience*, *29*(44), 13992–14003. https://doi.org/10.1523/JNEUROSCI.3577-09.2009

Brouwer, G. J., & Heeger, D. J. (2013). Categorical clustering of the neural representation of color. *Journal of Neuroscience*, *33*(39), 15454–15465. https://doi.org/10.1523/JNEUROSCI.2472-13.2013

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. MIT Press.

Crognale, M. A. (2002). Development, maturation, and aging of chromatic visual pathways: VEP results. *Journal of Vision*, *2*(6), 438–450. https://doi.org/10.1167/2.6.2

Crognale, M. A., Switkes, E., Rabin, J., Schneck, M. E., Hægerström-Portnoy, G., & Adams, A. J. (1993). Application of the spatiochromatic visual evoked potential to detection of congenital and acquired color-vision deficiencies. *Journal of the Optical Society of America A*, *10*(8), 1818–1825. https://doi.org/10.1364/josaa.10.001818

Forder, L., Bosten, J., He, X., & Franklin, A. (2017). A neural signature of the unique hues. *Scientific Reports*, *7*, 1–8. https://doi.org/10.1038/srep42364

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, *115*(1), 168–177. https://doi.org/10.1152/jn.00860.2015

Fukuda, K., Mance, I., & Vogel, E. K. (2015). Power modulation and event-related slow wave provide dissociable correlates of visual working memory. *Journal of Neuroscience*, *35*(41), 14009–14016. https://doi.org/10.1523/JNEUROSCI.5003-14.2015

Kim, I., Hong, S. W., Shevell, S. K., & Shim, W. M. (2020). Neural representations of perceptual color experience in the human ventral visual pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(23), 13145–13150. https://doi.org/10.1073/pnas.1911041117

Klistorner, A., Crewther, D. P., & Crewther, S. G. (1998). Temporal analysis of the chromatic flash VEP - Separate colour and luminance contrast components. *Vision Research*, *38*(24), 3979–4000. https://doi.org/10.1016/S0042-6989(97)00394-5

Kulikowski, J. J., Robson, A. G., & Mckeefry, D. J. (1996). Specificity and selectivity of chromatic visual evoked potentials. *Vision Research*, *36*(21), 3397–3401. https://doi.org/10.1016/0042-6989(96)00055-7

Lee, B., Martin, P., & Valberg, A. (1988). The physiological basis of heterochromatic flicker photometry. *Journal of Physiology*, *404*, 323–347.

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Murray, I. J., Parry, N. R. A., Carden, D., & Kulikowski, J. J. (1987). Human visual evoked potentials to chromatic and achromatic gratings. *Clinical Vision Sciences*, *1*(3), 231–244.

Nunez, V., Shapley, R. M., & Gordon, J. (2018). Cortical double-opponent cells in color perception: Perceptual scaling and chromatic visual evoked potentials. *I-Perception*, *9*(1). https://doi.org/10.1177/2041669517752715

Paulus, W. M., Hömberg, V., Cunningham, K., Halliday, A. M., & Rohde, N. (1984). Colour and brightness components of foveal visual evoked potentials in man. *Electroencephalography and Clinical Neurophysiology*, *58*(2), 107–119. https://doi.org/10.1016/0013-4694(84)90023-3

Paulus, W. M., Hömberg, V., Cunningham, K., & Halliday, A. M. (1986). Colour and brightness coding in the central nervous system: Theoretical aspects and visual evoked potentials to homogeneous red and green stimuli. *Proceedings of the Royal Society of London - Biological Sciences*, *227*(1246), 53–66. https://doi.org/10.1098/rspb.1986.0009

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Rabin, J., Switkes, E., Crognale, M., Schneck, M. E., & Adams, A. J. (1994). Visual evoked potentials in three-dimensional color space:

Correlates of spatio-chromatic processing. *Vision Research*, *34*(20), 2657–2671. https://doi.org/10.1016/0042-6989(94)90222-4

Rasheed, S., & Marini, D. (2015). Classification of EEG signals produced by RGB colour stimuli. *Journal of Biomedical Engineering and Medical Imaging*, *2*(5), 56–69. https://doi.org/10.14738/jbemi.25.1566

Regan, D., & Spekreuse, H. (1974). Evoked potential indications of colour blindness. *Vision Research*, *14*(1), 89–95. https://doi.org/10.1016/0042-6989(74)90120-5

Sandhaeger, F., von Nicolai, C., Miller, E. K., & Siegel, M. (2019). Monkey EEG links neuronal color and motion information across species and scales. *eLife*, *8*, 1–21. https://doi.org/10.7554/eLife.45645

Shevell, S. K., & Martin, P. R. (2017). Color opponency: tutorial. *Journal of the Optical Society of America A*, *34*(7), 1099. https://doi.org/10.1364/josaa.34.001099

Siuda-Krzywicka, K., Boros, M., Bartolomeo, P., & Witzel, C. (2019). The biological bases of colour categorisation: From goldfish to the human brain. *Cortex*, *118*, 82–106. https://doi.org/10.1016/j.cortex.2019.04.010

Skiba, R. M., Duncan, C. S., & Crognale, M. A. (2014). The effects of luminance contribution from large fields to chromatic visual evoked potentials. *Vision Research*, *95*, 68–74. https://doi.org/10.1016/j.visres.2013.12.011

Sutterer, D. W., Foster, J. J., Serences, J. T., Vogel, E. K., & Awh, E. (2019). Alpha-band oscillations track the retrieval of precise spatial representations from long-term memory. *Journal of Neurophysiology*, *122*(2), 539–551. https://doi.org/10.1152/jn.00268.2019

van Moorselaar, D., Foster, J. J., Sutterer, D. W., Theeuwes, J., Olivers, C. N. L., & Awh, E. (2017). Spatially selective alpha oscillations reveal moment-by-moment trade-offs between working memory and attention. *Journal of Cognitive Neuroscience*, *30*(2), 256–266. https://doi.org/10.1162/jocn

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. https://doi.org/10.1038/nn.4546

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

FIGURE S1 Classification of chromaticity, luminance, and joint chromaticity and luminance from total power in Experiment 2 and Experiments 1 and 2 combined. (a) Decoding presented chromaticity from the topographic distribution of total power across a range of frequencies (4–50 Hz) in Experiment 2.* (b) Decoding the presented luminance from the topographic distribution of total power across a range of frequencies (4–50 Hz) in Experiment 2.* (c) Decoding the specific luminance and chromaticity of stimuli from the topographic distribution of total power across a range of frequencies (4–50 Hz) in Experiment 2.* (d) Decoding presented chromaticity from the topographic distribution of total power across a range of frequencies (4 – 50 Hz) for Experiments 1 and 2 combined.* (e) Decoding the presented luminance from the topographic distribution of total power across a range of frequencies (4–50 Hz) for Experiments 1 and 2 combined.* (f) Decoding the specific luminance and chromaticity of stimuli from the topographic distribution of total power across a range of frequencies (4–50 Hz) in Experiments 1 and 2 combined.* *Time frequency points where decoding was not reliably above chance as determined by a cluster corrected permutation test ($p < 0.05$) were set to dark blue (the lowest value of the color scale)

FIGURE S2 ERP voltage topography by category for Experiment 2 and Experiments 1 and 2 combined. (a) Average ERP voltage topography (100–500 ms) by category for Experiment 2 (b) Average ERP voltage topography (100–500 ms) by category for Experiments 1 and 2 combined.

FIGURE S3 ERPs at midline electrodes by category for Experiment 2 and Experiments 1 and 2 combined. (a) ERP voltage at midline electrodes for Experiment 2. (b) ERP voltage at midline electrodes for Experiments 1 and 2 combined. Line color reflects chromaticity, dotted lines reflect low luminance, and solid lines reflect medium luminance

How to cite this article: Sutterer DW, Coia AJ, Sun V, Shevell SK, Awh E. Decoding chromaticity and luminance from patterns of EEG activity. *Psychophysiology*. 2021;58:e13779. https://doi.org/10.1111/psyp.13779